UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION

1b. RESTRICTIVE MARKINGS

AD-A216 456

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release;
distribution unlimited.

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 89-1719

6a. NAME OF PERFORMING ORGANIZATION
University of Connecticut

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS *(City, State, and ZIP Code)*
Department of Mathematics
U-9 Room 111
196 Auditorium Road, Storrs, CT 06268

7b. ADDRESS *(City, State, and ZIP Code)*
Building 410
Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL *(If applicable)*
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-86-0145

8c. ADDRESS *(City, State, and ZIP Code)*
Building 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A9	

11. TITLE *(Include Security Classification)*
INVERSE SCATTERING AND TOMOGRAPHY

12. PERSONAL AUTHOR(S)
Professor W. R. Madych

13a. TYPE OF REPORT
FINAL

13b. TIME COVERED
FROM 1 Jun 86 TO 31 May89

14. DATE OF REPORT *(Year, Month, Day)*

15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Conditions are prescribed for a function g which are sufficient to ensure that it is the Randon transform of a continuous function f on $\mathbb{R}^n$ such that $f(x) = o(|x|^{-n-k-1})$ as $|x| \to \infty$. Roughly speaking, these criteria involve smoothness and the classical polynomial consistency conditons up to order k on g. In particular, the result implies Helgason's Schwartz theorem for the Radon transform [Acta Math. 113, 1965].

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
DR. ARJE NACHMAN

22b. TELEPHONE *(Include Area Code)*
(202) 767- 4939

22c. OFFICE SYMBOL
NM

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

89 12 20 013 UNCLASSIFIED

This report includes these types

Contents

INVERSE SCATTERING AND TOMOGRAPHY

Final Technical Report for AFOSR-86-0145.
November 27, 1989
PI - W. R. Madych

The following is a list of work completed while the principal investigator was partially supported by this grant.
The following are either published or accepted for publication:

1. A Range Theorem for the Radon Transform, *Proc. Amer. Math. Soc.*, 104 (1988), 79-85, with D. C. Solmon.

2. Multivariate Interpolation and Conditionally Positive Definite Function, *J. Approx. Theory and Its Appl.*, 4 (1988), 77-89, with S. A. Nelson

3. Polyharmonic Cardinal Splines, to appear in *J. Approx. Theory*, with S. A. Nelson, 16pp.

4. Solutions of underdetermined systems of linear equations, to appear in *Proc. 1988 AMS-IMS-SIAM Joint Conference on Spatial Statistics and Imaging*, 16 pp.

5. Multivariate interpolation and conditionally positive definite functions II, to appear in *Math. Computation*, with S. A. Nelson, 19pp.

6. Polyharmonic Cardinal Splines: a Minimization Property, *J. Approx. Theory*, with S. A. Nelson, 16pp.

7. Error bounds for multiquadric interpolation, to appear in *Approximation Theory VI*, with S. A. Nelson, 4pp.

8. On the correctness of the problem of inverting the finite Hilbert transform in certain aeroelastic models, to appear in *J. Integral Equations and Applications*, 5pp.

9. Cardinal interpolation with polyharmonic splines, to appear in *Proc. 1989 Oberwolfach Conf. on Multivariate Approx.*, 8pp.

The following are preprints submitted for publication and technical reports:

1. Hilbert spaces for estimators, BRC/Math-TR-88-5, 12pp.

2. Translation invariant multiscale analysis, 8pp.

A copy of each of the above works is included in the enclosed volume.

This grant also provided partial support for one of my students, Dr. W. Yu, for the preparation of his PhD. thesis entitled *Inverse Problems in Partial Differential Equations*. A copy of this work is enclosed.

This grant provided at least part of the cost of participation in each of the following conferences.

1. August '86 – Oberwolfach, Germany. Participated in conference on the Radon Transformation and its applications. Gave talk on the limited angle problem and related material.

2. January '87 – Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN. Participated in period of concentration on inverse problems. Gave talk on the limited angle problem and related material.

3. April '87 – American Mathematical Society Meeting, Newark, N.J. Gave talk in special session on computational mathematics. Title of talk: K-harmonic splines.

4. July '87 – SIAM Conference in Albany, NY. Gave talk (contributed) on K-harmonic splines.

5. June '88 – AMS-IMS-SIAM Joint conference on spatial statistics and imaging, Brunswick, ME. Talk on solutions of underdetermined systems in image processing.

6. August '88 – NATO conference on numerical linear algebra. Talk on solutions of underdetermined systems.

7. January '89 – Approximation theory conference in College Station, TX. Talk on error bound for multiquadric interpolation.

8. February '89 – Oberwolfach, Germany. Participated in conference on multivariate approximation and applications. Gave talk on k-harmonic cardinal splines.

9. May '89 – Participated in the international conference on Wavelets in Marseille.

10. June '89 – Analysis conference in El Escorial. Gave talk on splines and entire functions.

11. June '89 – AMS summer institute on integral geometry, Arcata, CA. Gave talk on computed tomography.

12. July '89 – NATO ASI in Tuscany, Italy. Gave talk on splines and entire functions.

13. August '89 – International Workshop on Multivariate Approximation and Interpolation, Duisburg, W. Germany. Gave talk on multiscale analysis and polyharmonic splines.

A RANGE THEOREM FOR THE RADON TRANSFORM

by

W. R. Madych[1]

Department of Mathematics
University of Connecticut
Storrs, Connecticut 06268

D. C. Solmon[2]

Department of Mathematics
Oregon State University
Corvallis, Oregon 97331

ABSTRACT

Conditions are prescribed for a function g which are sufficient to ensure that it is the Radon transform of a continuous function f on $\mathbb{R}^n$ such that $f(x) = O(|x|^{-n-k-1})$ as $|x| \longrightarrow \infty$. Roughly speaking, these criteria involve smoothness and the classical polynomial consistency conditions up to order k on g. In particular, the result implies Helgason's Schwartz theorem for the Radon transform [Acta Math. 113, 1965].

Mathematics Subject Classification - 44A15, 26B40

Key Words - Radon transform, ploynomial consistency condition, asymptotic behavior

[1]Supported by Air Force Office of Scientific Research Grant No. AFOSR-86-1045.

[2]Supported by NSF Grant No. DMS-8602300 and Consiglio Nazionale Delle Ricerche, Firenze, Italia.

1. Introduction

The Radon transform of an integrable function f on $\mathbb{R}^n$ is defined by

$$R_\theta f(t) = \int_{\langle x,\theta\rangle = t} f(x)dx .$$

Here θ is a direction, i.e., a point on the unit sphere S^{n-1}, t is a real number and the integral is over the hyperplane orthogonal to θ and passing a directed distance t from the origin. We address the questions of when a given function g on $S^{n-1}\times\mathbb{R}$ is the Radon transform of a function f and what regularity and decay conditions on f can be deduced from those of g. The fundamental result on this quesiton is due to Helgason [3]. His theorem says that an even function g in the Schwartz space $\mathcal{S}(S^{n-1}\times\mathbb{R})$ is the Radon transform of a function f in the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ if and only if

$$\int_{-\infty}^{\infty} g(\theta,t)t^j dt = P_j(\theta), \quad j = 0,1,... \tag{1.1}$$

is representable as a homogeneous polynomial of degree j in θ. He also showed that under the above hypotheses f has compact support if and only if g does and that the convex hull of the support of f is determined by the support of g. The first result is called the Schwartz theorem for the Radon transform and the second the Paley-Wiener theorem for the Radon transform [4]. An L^2

version of the Paley-Wiener theorem for the Radon transform was obtained by Lax and Phillips [6]. Smith et. al. [9] showed that any even function in the Sobolev space $H^s(S^{n-1}\times\mathbb{R})$, $s = (n-1)/2$, is in the range of the closure $\bar{R}$ of the Radon transform as an unbounded operator on $L^2(\mathbb{R}^n)$, but that $\bar{R}$ is not necessarily defined by an absolutely convergent integral. Recently, Solmon [10] showed that any even g in $S(S^{n-1}\times\mathbb{R})$ is the Radon transform of a C^∞ function f such that $f(x) = O(|x|^{-n})$ as $|x| \longrightarrow \infty$, and that $f(x) = O(|x|^{-n-k-1})$ if and only if (1.1) holds for $j = 0,1,\ldots,k$.

In this paper we show that a sufficiently smooth even function on $S^{n-1}\times\mathbb{R}$ which together with a finite number of derivatives decays sufficiently fast at ∞ is the Radon transform of a continuous function f that is $O(|x|^{-n})$ as $|x| \longrightarrow \infty$, and show how satisfying a finite number of the conditions (1.1) influences the behavior of f at ∞ . (Of course precise conditions on the smoothness and decay are given.) In particular, the result given here implies the Schwartz theorem of Helgason and the recent extension by Solmon. Our proof is based on a result on the asymptotic behavior of the Fourier transform due to Madych [7], rather than the Radon inversion formula as in [10]. This approach is independent of dimension, shorter and simpler than that in [10]. Moreover, the theorem given here is more general.

2. Operators and Formulas

Let θ be a point on the unit sphere S^{n-1} in $\mathbb{R}^n$, $n \geq 2$, and t be a real number. The Radon transform of an integrable function f on $\mathbb{R}^n$ in the direction θ at the point t is the function

$$(2.1) \qquad Rf(\theta,t) = R_\theta f(t) = \int_{\langle x,\theta\rangle=t} f(x)dx ,$$

where $\langle\ ,\ \rangle$ denotes the usual Euclidean inner product on $\mathbb{R}^n$ and dx represents integration with respect to $n-1$ dimensional Lebesgue measure on the hyperplane $\langle x,\theta\rangle = t$.

The Fourier transform of an integrable function f on $\mathbb{R}^n$ is defined by

$$\hat{f}(\xi) = (2\pi)^{-n//2} \int f(x)e^{-i\langle x,\xi\rangle}dx ,$$

where the integral is taken over all of $\mathbb{R}^n$. Fixing θ in (2.1) and taking the one dimensional Fourier transform gives the well known formula

$$(2.2) \qquad (R_\theta f)\hat{}(\tau) = (2\pi)^{(n-1)/2}\hat{f}(\tau\theta) .$$

The operator Λ is defined in terms of Fourier transforms by

$$(2.3) \qquad (\Lambda f)\hat{}(\xi) = |\xi|\hat{f}(\xi) ,$$

so that $\Lambda^2 = -\Delta$, where Δ is the Laplacean. We shall also use

the Calderon-Zygmund representation of Λ, [1],

$$(2.4) \qquad \Lambda = \sum_{j=1}^{n} \mathcal{H}_j D_j ,$$

where D_j denotes partial differentiation with respect to the j-th coordinate and $\mathcal{H}_j$ is the singular integral operator defined by convolution with the kernel $c_n x_j/|x|^{n+1}$, c_n a constant depending only on n ; i.e., $\mathcal{H}_j$ is the j-th Riesz transform.

The dual Radon transform is the formal adjoint of the Radon transform and is defined on functions on $S^{n-1}\times\mathbb{R}$ by

$$(2.5) \qquad R^t g(x) = \int_{S^{n-1}} g(\theta,\langle x.\theta\rangle)d\theta .$$

The Radon transform, dual Radon transform, and Calderon-Zygmund operator are related by the Radon inversion formula

$$(2.6) \qquad f = 2^{-1}(2\pi)^{1-n}\Lambda^{n-1}R^t Rf .$$

Remark 2.7. If $f \in L^p(\mathbb{R}^n)$, $1 < p < n/(n-1)$, then for almost every θ, $R_\theta f(t)$ exists for almost every t and for a.e. θ (2.2) is valid for a.e. τ . Also (2.6) holds almost everywhere on $\mathbb{R}^n$. See [10].

We use the standard multi-index notation. Thus if ν is a multi-index i.e. an n-tuple of nonnegative integers, $x \in \mathbb{R}^n$, and $D = (D_1,\dots,D_n)$, then $x^\nu = x_1^{\nu_1}\dots x_n^{\nu_n}$, D^ν is the differential

operator $D_1^{\nu_1}\cdots D_n^{\nu_n}$, and $|\nu| = \nu_1+\cdots+\nu_n$. Also, D_θ^j denotes a j-th order (tangential) differential operator on the sphere S^{n-1} with smooth coefficients and $g^{(j)}(\theta,t)$ denotes the j-th order partial derivative of $g(\theta,t)$ with respect to t .

A function f on $\mathbb{R}^n$ will be said to be $O(|x|^{-m})$ as $|x| \longrightarrow \infty$ if there exist constants $C, M > 0$ such that $|f(x)| \le C|x|^{-m}$ whenever $|x| \ge M$. The analogous definition is taken for the statements $f(x) = O(|x|^{-m})$ as $|x| \longrightarrow 0$, and $g(\theta,t) = O(|t|^{-m})$ as $|t| \longrightarrow \infty$ or 0 .

The proof of the theorem on the range of the Radon transform is based on the following asymptotic result about Fourier transforms which is a straightforward generalization of [7, Proposition 5].

Lemma: Let f be a function on $\mathbb{R}^n$ such that $\hat{f}$ is integrable. Suppose that $\alpha > 0$, m is an integer satisfying $m > n + \alpha$, and that

(i) $|\xi|^{-\alpha}\hat{f}(\xi)$ is in $C^m(\mathbb{R}^n\setminus\{0\})$, and

(ii) $|\xi|^{|\nu|}D^\nu(|\xi|^{-\alpha}\,\hat{f}(\xi))$ is bounded on $\mathbb{R}^n$ for all ν such that $0 \le |\nu| \le m$.

Then f is continuous and $f(x) = O(|x|^{-n-\alpha})$ as $|x| \longrightarrow \infty$.

Proof: Since $\hat{f}$ is in $L^1(\mathbb{R}^n)$, f is continuous and bounded and it suffices to show that

$$|f(x)| \le C|x|^{-n-\alpha} \quad \text{if} \quad |x| > 1 . \tag{2.8}$$

To see (2.8) let $|x| > 1$, set $r = |x|$, $x' = x/r$, and write

$$f(x) = (2\pi)^{-n/2} \int |\xi|^{\alpha} h(\xi) e^{i\langle x,\xi\rangle} d\xi$$

$$= (2\pi)^{-n/2} r^{-n-\alpha} \int |\xi|^{\alpha} h(\xi/r) e^{i\langle x',\xi\rangle} d\xi \ ,$$

where $h(\xi) = |\xi|^{-\alpha} \hat{f}(\xi)$.

The last formula indicates that to verify (2.8) it suffices to show that the Fourier transform of $h_r(\xi) = |\xi|^{\alpha} h(\xi/r)$ is bounded on the unit sphere independent of $r > 1$.

To see this let φ be a function in $C^{\infty}(\mathbb{R}^n)$ such that $\varphi(\xi) = 1$ for $|\xi| < 1$ and $\varphi(\xi) = 0$ for $|\xi| > 2$ and write $h_r = h_0 + h_{\infty}$ where $h_0(\xi) = h_r(\xi)\varphi(\xi)$ and $h_{\infty}(\xi) = h_r(\xi)(1-\varphi(\xi))$. Since $\|h_0\|_1$ is bounded independent of r it follows that $\hat{h}_0(x)$ is bounded for all x and hence is certainly bounded on the unit sphere. To obtain a similar conclusion concerning h_{∞} write

$$D^{\nu} h_{\infty}(\xi) = A(\xi) + B(\xi) \ . \tag{2.9}$$

where $A(\xi) = \sum c_{\mu} |\xi|^{\alpha-|\mu|} D^{\nu-\mu} h(\xi/r) r^{-|\nu-\mu|} (1-\varphi(\xi))$, the c_{μ}'s are appropriate constants, and the sum is taken over all multi-indices μ such that $0 \le \mu \le \nu$. $B(\xi)$ is a similar expression except that each term contains a derivative of $(1-\varphi)$ of order one or higher. Observe that

$$|\xi|^{\alpha-|\mu|} D^{\nu-\mu} h(\xi/r) r^{-|\nu-\mu|} = |\xi|^{\alpha-|\nu|} \left[|\xi/r|^{|\nu|-|\mu|} D^{\nu-\mu} h(\xi/r) \right]$$

and the expression in paranthesis is bounded by virtue of statement (ii) in the hypothesis. Thus

$$|A(\xi)| \le C|\xi|^{\alpha-|\nu|}(1-\varphi(\xi)) \ . \tag{2.10}$$

A similar observation shows that B is supported in $1 \le |\xi| \le 2$ and is bounded. These facts concerning A and B imply that

$$|D^{\nu}h_{\infty}(\xi)| \le c|\xi|^{\alpha-|\nu|}(1-\varphi(2\xi)) \ . \tag{2.11}$$

Now, if $|\nu| = m$, recalling that m is greater than $n+\alpha$, it follows from (2.11) that $\|D^{\nu}h_{\infty}\|_1$ is bounded independent of r. Since this is true for all such ν, we may conclude that $||x|^m \hat{h}_{\infty}(x)|$ is bounded independent of r and hence $\hat{h}_{\infty}$ is bounded on the unit sphere. This implies the desired result. □

3. The Range Theorem and Corollaries

It is convenient to introduce the following definition.

Definition. *A function* g *on* $S^{n-1}\times\mathbb{R}$ *is said to be uniformly integrable if there exists an integrable function* h *on* $\mathbb{R}$ *such that* $|g(\theta,t)| \le h(t)$ *for all* θ *and* t.

The range theorem follows.

Theorem. *Let* $g \in C^r(S^{n-1}\times\mathbb{R})$, *where* $r \ge m = n+2+k$, $k \ge -1$. *Assume* $g(-\theta,-t) = g(\theta,t)$,

(3.1) $t^{k+1+q} D_\theta^j\, g^{(q)}(\theta,t)$ *is uniformly integrable for* $j = 0,1,\dots,m-q$, *and* $q = 0,1,\dots,n+1$; *and*

(3.2) $\int_{-\infty}^{\infty} g(\theta,t)t^\ell dt = P_\ell(\theta)$ *is representable as a homogeneous polynomial of degree* ℓ *for* $\ell = 0,1,\dots,k$.

Then there exists a function f *in* $C^s(R^n)$, ($s = r+1-n$, n *odd and* $s = r-n$, n *even*) , *such that* $f(x) = O(|x|^{-n-k-1})$ *as* $|x| \longrightarrow \infty$ *and* $Rf(\theta,t) = g(\theta,t)$ *for all* θ *and* t . *Moreover, for each* $x \in \mathbb{R}^n$ *the inversion formula*

$$f(x) = 2^{-1}(2\pi)^{1-n} \Lambda^{n-1} R^t g(x) \tag{3.3}$$

holds.

Remark. When $k = -1$, (3.2) is taken to be vacuous.

Proof. The continuity of g and (3.1) with $q = j = 0$ imply that g is uniformly integrable. Thus, the function $\hat{f}$ defined in polar coordinates on $\mathbb{R}^n$ by

$$\hat{f}(\tau\theta) = (2\pi)^{(1-n)/2} \hat{g}_\theta(\tau) = (2\pi)^{-n/2} \int_{-\infty}^{\infty} g(\theta,t)e^{-it\tau}dt \tag{3.4}$$

is bounded on $\mathbb{R}^n$ and continuous on $\mathbb{R}^n-\{0\}$. We show that $\hat{f}$ satisfies the hypotheses of the Lemma with $\alpha = k+1$. Leibnitz formula gives

$$(3.5) \qquad (d/dt)^{n+1}(t^m g) = \sum_{q=0}^{n+1} c_{m,k} t^{k+1+q} g^{(q)} ,$$

with the $c_{m,k}$ known constants. By (3.1) each term on the right hand side of (3.5) is uniformally integrable and hence so is the left hand side. Since $g^{(n+1)}$ is continuous, solving for $g^{(n+1)}$ in (3.5) shows that $g^{(n+1)}$ is uniformly integrable. Thus $\hat{g}$ is continuous on $S^{n-1}\times\mathbb{R}$ and uniformly $O(\tau^{-n-1})$ as $\tau \longrightarrow \infty$. This, (3.4), and the fact that $\hat{f}$ is bounded imply that $\hat{f}$ is integrable and hence its inverse Fourier transform f , is continuous.

We will show at the end of the proof that the polynomial condition (3.2) can be replaced by

$$(3.6) \qquad \int_{-\infty}^{\infty} g(\theta,t)t^{\ell} dt = 0 \ , \ \ell = 0,1,\ldots,k ,$$

without loss of generality. Accepting this for now, the uniform integrability assumption of (3.1) allows differentiation under the integral sign in (3.6) giving

$$\int_{-\infty}^{\infty} D_{\theta}^{j}\, g(\theta,t)t^{\ell}\, dt = 0 \ , \ \ell = 0,1,\ldots,k \ ; \quad j = 0,1,\ldots,m \ .$$

Since $t^{k+1}D_{\theta}^{j}\, g$ is uniformly integrable, the above and Taylor's theorem give that

$$(3.7) \qquad |(t^p D_{\theta}^{j}\, g)^{\wedge}(\tau)| = O(\tau^{k+1-p}) \quad \text{as} \quad \tau \longrightarrow 0 \ , \ \text{when} \quad 0 \le p \le k+1 \ .$$

Let $0 \le |\nu| \le m = n+2+k$. By (3.4), in polar coordinates $|\xi|^{|\nu|} D^{\nu}(|\xi|^{-k-1}\hat{f})$ is a sum of terms of the form $\tau^{\ell}(\partial/\partial\tau)^{\ell}(\tau^{-k-1}(D_{\theta}^{j}\, g)\hat{\ }(\tau\theta))$, $\ell = 0,1,\ldots,|\nu| \le m$, $j = 0,1,\ldots,|\nu|-\ell$. This, in turn, is a sum of terms of the form

$$(3.8) \qquad \tau^{-k-1+p}(t^{p}D_{\theta}^{j}\, g)\hat{\ }(\tau\theta), \quad 0 \le p \le \ell \le |\nu| \le m,$$

where the Fourier transform is in the sense of distributions when $p > k+1$. If $0 \le p \le k+1$, (3.7) and the uniform integrability assumption (3.1) show that terms of the form (3.8) are continuous on $\mathbb{R}^{n}-\{0\}$ and bounded independent of θ. Suppose $m \ge |\nu| \ge \ell \ge p \ge k+1$. Leibnitz formula shows that $(d/dt)^{p-k-1}(t^{p}D_{\theta}^{j}g)$ is a sum of terms of the form $t^{k+1+q}D_{\theta}^{j}\, g^{(q)}$, $q = 0,1,\ldots,p-k-1$. Each such term is uniformly integrable by (3.1). Thus $(t^{p}D_{\theta}^{j}\, g)\hat{\ }(\tau\theta)$ is continuous on $\tau > 0$ and uniformly $O(\tau^{k+1-p})$ as $\tau \longrightarrow \infty$ and 0. Hence (3.8) is bounded in this case also. The Lemma now implies that $f(x) = O(|x|^{-n-k-1})$ as $|x| \longrightarrow \infty$.

Since f is continuous and $O(|x|^{-n})$ as $|x| \longrightarrow \infty$, Rf is continuous also. Moreover $f \in L^{p}(\mathbb{R}^{n})$ for all $p > 1$. By Remark (2.7), (2.2) holds a.e. This, together with (3.4), shows that $Rf = g$ a.e. and thus, by continuity, everywhere. Again Remark (2.7) gives that the inversion formula (2.6), and hence (3.3), holds a.e. To show that it holds everywhere it suffices to show that the right hand side of (3.3) is continuous. If n is odd, then (2.3), (2.5) and the fact that $g \in C^{r}(S^{n-1}\times\mathbb{R})$ show that the right hand side of (3.3), and hence f, is C^{r+1-n}. If n is even, then write

$\Lambda^{n-1} = \Lambda(\Delta^{(n-2)/2})$ and use the Calderon-Zygmund representation of Λ, (2.4). This, together with the fact that the Riesz transforms, $\mathcal{R}_j$, reduce differentiability by at most ϵ for any $\epsilon > 0$, shows that the right hand side of (3.3), and hence f, is C^{r-n}. This completes the proof of the Theorem with the exception of showing that, without loss of generality, (3.2) can be replaced by (3.6). To see this, note that by [10, Lemma 7.4], there exists a C^∞ function h vanishing outside of the unit ball in $\mathbb{R}^n$ such that

$$\int_{-\infty}^{\infty} R_\theta h(t)t^\ell \, dt = P_\ell(\theta) \ , \ \ell = 0,1,\ldots,k \ ,$$

where the P_ℓ are the same polynomials as those that appear in (3.2). Replacing g by $g - Rh$ does not effect (3.1) but replaces (3.2) by (3.6), and f by $f - h$ which has no effect on the smoothness of f or its rate of decay at ∞. The proof is complete.

<u>**Remark**</u>. The proof actually shows that when n is even then all partial derivatives of f or order $r - n$ satisfy a Hölder condition of order α for any $\alpha < 1$. Also, it follows from [9, Theorem 12.6] that the classical Radon inversion formula $f = 2^{-1}(2\pi)^{-n}R^t\Lambda^{n-1}g$ is valid everywhere.

We now give a few corollaries of the Theorem.

The Schwartz space $\mathcal{S}(S^{n-1}\times\mathbb{R})$ consists of those C^∞ functions on $S^{n-1}\times\mathbb{R}$ which, together with their derivatives of all orders,

decay at infinity faster than the reciprocal of any polynomial in t.

Corollary 1. [10]. *Let* $g \in S(S^{n-1} \times \mathbb{R})$ *and suppose that* $g(-\theta, -t) = g(\theta, t)$. *Then*

a) $g = Rf$ *for some* C^∞ *function* f *such that* $f(x) = O(|x|^{-n})$ *as* $|x| \longrightarrow \infty$.

b) $f(x) = O(|x|^{-n-k-1})$ *if and only if* (3.2) *holds for* $\ell = 0,1,\ldots,k$.

c) *If* $f(x) = O(|x|^{-m})$ *then for all* $\nu \in N^n$, $D^\nu f(x) = O(|x|^{-m-|\nu|})$.

Proof. When $g \in S(S^{n-1} \times \mathbb{R})$, (3.1) is satisfied for all choices of j, q and k. So a) and b) are immediate consequences of the Theorem. To see that c) holds observe from (2.2) that $R_\theta D^\nu f = \theta^\nu (R_\theta f)^{(|\nu|)}$ and by integration by parts that if $g = Rf$ satisfies (3.2) for $\ell = 0,1,\ldots,k$, then $\theta^\nu g^{(|\nu|)}$ satisfies (3.2) for $\ell = 0,1,\ldots,|\nu|+k$, and apply b).

Corollary 2. *Let* $g \in C^{n+1}(S^{n-1} \times \mathbb{R})$ *and suppose that* $g(-\theta,-t)=g(\theta,t)$. *Assume that* (3.1) *is satisfied with* $k = -1$. *Then the inversion formula for* R^t,

$$g = 2^{-1}(2\pi)^{1-n} R \Lambda^{n-1} R^t g, \tag{3.9}$$

holds everywhere.

Proof. By the Theorem $Rf = g$ everywhere with f given by (3.3).

Remark. Inversion formulas for R^t have been given earlier for functions in $S(S^{n-1}\times\mathbb{R})$. Helgason [3] gave an inverison formula when (3.6) is satisfied for all nonnegative integers ℓ. Gonzalez [2] and Hertle [5] proved that (3.9) holds for n odd and Solmon [10] showed it holds in all dimensions.

The last corollary is a sort of Tauberian theorem.

Corollary 3. *Let* $g \in C^{n+2}(S^{n-1}\times\mathbb{R})$ *be such that* $g(-\theta,-t) = g(\theta,t)$. *Suppose that* (3.1) *is satisfied for* $k = 0$. *Then* g *is the Radon transform of an integrable function* f *if and only if*

$$\int_{-\infty}^{\infty} g(\theta,t)dt = c \,, \tag{3.10}$$

where c *is a constant, independent of* θ.

Proof. Since the hypotheses of the Theorem are satisfied with $k = 0$, g is the Radon transform of a continuous function that is $O(|x|^{-n-1})$ as $|x| \to \infty$, and hence of an integrable function. So the condition is sufficient. The necessity of the condition is an immediate consequence of Fubini's theorem.

Remark. Peters [8] has shown that the Radon transform of the function $f(x) = \sin(|x|^2)/(|x|^2)$, in the plane, is $O(|t|^{-2})$ as

$t \longrightarrow \infty$. Since f is a radial function its Radon transform is independent of θ and hence so is $\int_{-\infty}^{\infty} Rf(\theta,t)dt$. Nevertheless f is not integrable on $\mathbb{R}^2$. Hence condition (3.10) is not sufficient by itself to imply the integrability of f.

REFERENCES

1. Calderon, A. P. and Zygmund, A., Singular integral operators and differential equations, Amer. J. Math. *79*(1957), 901-921.

2. Gonzalez, F., Radon transforms on Grassmann manifolds, Ph.D. Thesis M.I.T., (1984).

3. Helgason, S., The Radon transform on Euclidean spaces, compact two point homogeneous spaces and Grassmann manifolds, Acta Math. *113*(1965), 153-180.

4. ____________, The Radon Transform, Boston: Birkhauser, (1980).

5. Hertle, A., Habilitationsschrift, U. Mainz (1986), (to appear).

6. Lax, P. D. and Phillips, R. S., The Paley-Wiener Theorem for the Radon Transform, Comm. on Pure and Applied Math. XXIII, (1970), 409-424.

7. Madych, W.R., On Littlewood-Paley functions, Studia Math., *50*(1974), 43-63.

8. Peters, J., A Tauberian theorem for the Radon transform, Houston J. Math. *8*(1982), 565-574.

9. Smith, K. T., Solmon, D. C., and Wagner, S. L., Practical and mathematical aspects of the problem of reconstructing objects from radiographs, Bull. A.M.S. *83*(1977), 1227-1270.

10. Solmon, D.C., Asymptotic formulas for the dual Radon transform and applications, M. Zeitschrift, 195-3(1987), 321-343.

Multivariate Interpolation and

Conditionally Positive Definite Functions

by

W. R. Madych [1]
Department of Mathematics, U-9
University of Connecticut
Storrs, CT. 06268

and

S. A. Nelson [1]
Department of Mathematics
Iowa State University
Ames, IA. 50011

Abstract: A theory of multivariate splines is presented which includes "Multiquadric" surfaces. It is essentially an extension of Duchon's theory, but this development completely avoids the use of Fourier analysis. More precisely we consider the problem of interpolating numerical data at a finite number of points in the plane or, more generally, in n space by interpolants which are essentially linear combinations of translates of a prescribed function h. A semi-norm, minimized by such interpolants is shown to exist if and only if h is conditionally positive definite. Pointwise error estimates are obtained in terms of this semi-norm.

1. Both authors were partially supported by a grant from Air Force Office of Scientific Research, AFOSR - 86 - 0145.

1. Introduction

A variety of methods exist for fitting an interpolating surface through data $v_1, \ldots, v_N$ given at a scattered set of points $x_1, \ldots, x_N$ in the plane $\mathbb{R}^2$ (or higher dimensional space $\mathbb{R}^n$). For a recent survey see [7]. Two methods that performed quite well in comparative numerical tests [6] are: Thin Plate Splines (TPS) [8] and Multiquadric Surfaces (MQS) [9]. A rather complete theory, applicable to TPS, but not MQS, has been developed by Duchon [3] - [5]; also see [14]. Theoretical results about MQS are more recent and less complete, [7], [12], [13].

In [12] an extension of Duchon's theory was developed which includes MQS. That development relies heavily on Fourier analysis. Here we present refinements of those results which include error estimates and avoid the use of Fourier transforms.

In this development certain features of Duchon's theory are preserved. In particular the interpolants $s(x)$, $x \in \mathbb{R}^n$, can be characterized as being solutions of a variational problem in which a translation invariant quadratic functional is minimized subject to constraints $s(x_i) = v_i$. As in Duchon's theory the quadratic functional is associated with a Hilbert space norm and a reproducing kernel $K(x,y) = h(x-y)$. For TPS, $h(x) = |x|^2 \log|x|$ and for MQS, $h(x) = -\sqrt{\delta^2 + |x|^2}$. Here $|x|^2 = \sum_{k=1}^{n} (x_k)^2$.

Unlike Duchon's exposition, where the choice of norm takes precedence, the development here starts with the function h and uses it to define an appropriate norm. Another difference in our results is that we do not require any sort of dilation invariance.

The interpolants s are expressible in terms of translates of the function h. In the simplest case ($m=0$, below) $s(x) = \sum_{j=1}^{N} c_j\, h(x-x_j)$. In the general case s includes a polynomial of degree less than some fixed integer m and the coefficients c_j are restricted by the requirement that $\sum_{j=1}^{N} c_j\, p(x_j) = 0$ for all polynomials p of degree not exceeding m-1. Thus we have

$$(1.1) \qquad s(x) = \sum_{j=1}^{N} c_j\, h(x-x_j) + \sum_{|\alpha|<m} k_\alpha\, x^\alpha$$

where the constants c_j and k_α must satisfy

$$(1.1a) \qquad \sum_{j=1}^{N} c_j h(x_i-x_j) + \sum_{|\alpha|<m} k_\alpha\, x_i^\alpha = v_i, \quad i=1,\ldots,N$$

$$(1.1b) \qquad \sum_{j=1}^{N} c_j\, x_j^\alpha = 0\ , \qquad |\alpha|<m.$$

As usual, $\alpha = (\alpha_1,\ldots,\alpha_n)$ denotes an n-tuple of non-negative integers, $|\alpha| = \alpha_1+\ldots+\alpha_n$ and $x^\alpha = \prod_{k=1}^{n} (x_k)^{\alpha_k}$.

For the case of Thin Plate Splines, the equations (1.1) have direct physical interpretations. For that case, h is as mentioned above and m is 2. The c_j's are proportional to forces applied at the points x_j and the moment conditions (1.1b) are required for equilibrium, see [9]. Analogous interpretations can be made for cubic splines on the line; there n=1, $h(x) = |x|^3$ and m=2.

It is useful to record some elementary observations about (1.1). In matrix notation we have

$$\begin{aligned} Ac + Bk &= v \\ B^T c \quad &= 0 \end{aligned}$$

where A is NxN and B is Nxm' with $m' = \dim P_{m-1}(\mathbb{R}^n) = \frac{(m-1+n)!}{(m-1)!n!}$. Letting R_B denote the range of B we see that (1.1b) says that c is in $R_B^\perp$, where $R_B^\perp = \left\{ c \in \mathbb{C}^N : \overline{b^T} c = 0 \text{ for all } b \in R_B \right\}$. Consequently, (1.1) will have a solution if and only if v is in $A(R_B^\perp) + R_B$. Now $\mathbb{C}^N = A(R_B^\perp) + R_B$ iff $Ac \notin R_B$ for all $c \in R_B^\perp \sim \{0\}$. Furthermore, $Ac \notin R_B$ $(=R_B^{\perp\perp})$ iff $\overline{w^T} Ac \neq 0$ for some $w \in R_B^\perp$. Thus, (1.1) <u>can be solved for arbitrary</u> v <u>if and only if for every</u> $c \in R_B^\perp \sim \{0\}$ <u>there is a</u> $w \in R_B^\perp$ <u>such that</u> $\overline{w^T} Ac \neq 0$. In addition we note that $\mathbb{C}^N = A(R_B^\perp) + R_B$ requires $\dim A(R_B^\perp) = \dim R_B^\perp$. Hence, <u>if</u> (1.1) <u>can be solved for arbitrary</u> v <u>then</u> c <u>in</u> (1.1) <u>is unique</u>.

A condition that is sufficient to make (1.1) solvable for all v is: $\overline{c^T} Ac > 0$ for all $c \in R_B^\perp \sim \{0\}$. This condition can be

restated as: $\sum_{i,j=1}^{N} \overline{c_i}\, h(x_i - x_j)\, c_j > 0$ for all $c \neq 0$ satisfying (1.1b).

The last inequality offers motivation for (1.2) in the following definition. However, the true motivation for (1.2) is that it is needed (see equation (1.3) below) in order to obtain our variational setting for the interpolation problem (1.1).

<u>Definition</u>: If h is a continuous function from $\mathbb{R}^n$ to $\mathbb{C}$ then h is <u>conditionally positive definite</u> of order m, provided that for every positive integer N,

$$\sum_{i,j=1}^{N} \bar{c}_i\, c_j\, h(x_i - x_j) \geq 0 \tag{1.2}$$

holds whenever $x_1, \ldots, x_N$ are distinct points in $\mathbb{R}^n$ and $c_1, \ldots, c_N$ are complex numbers that satisfy (1.1b). We define $Q_m(\mathbb{R}^n)$ to be the class of conditionally positive definite functions of order m on $\mathbb{R}^n$.

In the case m=0 the above definition reduces to the usual definition of a positive definite function. Thus, by Bochner's theorem, $Q_0(\mathbb{R}^n)$ is the set whose elements are Fourier transforms of finite positive Borel measures. For m>0, an analogous characterization of the functions in $Q_m(\mathbb{R}^n)$ is given in [8, Chapter II, Section 4.4].

It is not always easy to determine whether a given continuous function h is conditionally positive definite. Often the most direct method is to calculate the (distributional) Fourier

transform of h and check whether it is positive. This was done for several examples, including $h(x) = -\sqrt{\delta^2+|x|^2}$, in [12].

In the case that $h(x)$ is a radial function the task of determining whether it is conditionally positive definite can often be considerably simplified by applying Micchelli's generalization [13] of Schoenberg's theorem [14] concerning completely monotone functions. This result can be stated as follows: If $F(t)$ is real valued and continuous for $t \geq 0$ then $h(x) = F(|x|^2)$ is in $Q_m(\mathbb{R}^n)$ for all n if and only if $(-1)^{m+\ell} F^{(m+\ell)}(t) \geq 0$ for all $t \geq 0$ and all $\ell = 0,1,2,\ldots$. Application of this theorem with $F(t) = t \log t$, $m = 2$, and $F(t) = -\sqrt{\delta^2+t}$, $m = 1$, provides an easy verification of the conditional positive definiteness of $h(x) = |x|^2\log|x|$ and $h(x) = -\sqrt{\delta^2+|x|^2}$.

We now turn our attention to the variational setting for the interpolation problem. This setting is the space C_h described by Theorem 1.1 below. First we recall some notation.

Let $P_{m-1}(\mathbb{R}^n)$ denote the space consisting of polynomials on $\mathbb{R}^n$ of degree not exceeding $m-1$. We define $P_{m-1}^{\perp}(\mathbb{R}^n)$ to be space of all measures ν on $\mathbb{R}^n$ that have support consisting of a finite number of points and satisfy $\nu(p) = 0$ for all p in $P_{m-1}(\mathbb{R}^n)$.

Note that $\nu = \sum_{i=1}^{N} c_i \, \delta_{x_i}$ is in $P_{m-1}^{\perp}$ if and only if (1.1b) is satisfied. Here δ_x is the measure corresponding to a unit mass

at the point x, so $\delta_x(f) = f(x)$. In Theorem 1.1 $\nu * h$ denotes the usual convolution product. For example, $\delta_y * h(x) = h(x-y)$.

Theorem 1.1: Given $m \geq 0$ and h in $Q_m(\mathbb{R}^n)$ there is a subspace $C_h(\mathbb{R}^n)$ of $C(\mathbb{R}^n)$, with a semi-inner product $(\cdot,\cdot)_h$, such that

(a) the null-space of the semi- norm is $P_{m-1}(\mathbb{R}^n)$,

(b) C_h/P_{m-1} is a Hilbert space,

(c) if ν is in $P_{m-1}^{\perp}(\mathbb{R}^n)$ then $\nu * h$ is in $C_h(\mathbb{R}^n)$ and

$(\nu * h, f)_h = \nu(\bar{f})$ for all f in $C_h(\mathbb{R}^n)$

Furthermore, these properties uniquely determine $C_h(\mathbb{R}^n)$ and $(\cdot,\cdot)_h$.

The proof of Theroem 1.1 is carried out in section 3. It includes an explicit construction of the subspace C_h which does not involve Fourier analysis. On the other hand, except for certain special examples (see below), C_h is rather awkward to describe without this basic tool.

The results in Section 2 show that C_h contains interpolants of the data $v_1,\ldots,v_N$ at the scattered set of points $x_1,\ldots,x_N$. Moreover it is also shown that the interpolant of the form (1.1) is in C_h with minimum C_h norm.

The connection with reproducing kernel Hilbert spaces can be seen as follows. If m=0, we can take $\nu = \delta_x$ in part (c) of Theorem 1.1, to obtain

$$f(x) = \overline{(\delta_x * h, f)_h} = (f, \delta_x * h)_h$$

Thus, $K(x,y) = h(x-y)$ is a reproducing kernel for the space

C_h if m=0. For background on reproducing kernels and positive definiteness we recommend [2]. An extended notion of reproducing kernel is defined in [4] and [16]; in essence, Theorem 1.1 says that $K(x,y) = h(x-y)$ is such a kernel.

The necessity of (1.2) can be seen by taking $\nu = \sum_{i=1}^{N} c_i \, \delta_{x_i}$ and $f = \nu * h$ in (c) to get

$$(1.3) \quad (\nu * h, \nu * h)_h = \sum_{i=1}^{N} c_i \, \overline{\nu * h(x_i)} = \sum_{i,j=1}^{N} c_i \, \overline{c_j} \quad \overline{h(x_i - x_j)} \, .$$

We conclude this section with several examples of h and corresponding C_h.

A. $n=1$, $m=0$, $h(x) = e^{-|x|}$, $C_h = \{f \in AC \cap L^2(\mathbb{R}) : f' \in L^2(\mathbb{R})\}$,

$$(f,f)_h = \frac{1}{2} \int_{\mathbb{R}} \left\{ |f(x)|^2 + |f'(x)|^2 \right\} dx.$$

B. $n=1$, $m=1$, $h(x) = -|x|$, $C_h = \{f \in AC : f' \in L^2(\mathbb{R})\}$,

$$(f,f)_h = \frac{1}{2} \int_{\mathbb{R}} |f'(x)|^2 \, dx.$$

C. Consider the cases included in Duchon's theory. Example B, but not A, is such a case. Other examples include interpolants in $\mathbb{R}^n$ which satisfy $\Delta^m s(x) = 0$ whenever x is in $\mathbb{R}^n \backslash \{x_1, \ldots, x_N\}$ and which minimize the functional $(f,f)_h = \int_{\mathbb{R}^n} |\Delta^k f(x)|^2 dx$, where

Δ^m denotes the m-th iterate of the Laplace operator, m must be an integer greater than n/2, k=m/2, and Δ^k is suitably interpreted if k is not an integer. For details see [4]. Here $h(x) = c|x|^{m-n}$ or $c|x|^{m-n} \log|x|$ depending on whether n is odd or even; c is a normalizing constant depending only on m and n.

D. As mentioned earlier, the function $h(x) = -\sqrt{\delta^2+|x|^2}$ is in $Q_1(\mathbb{R}^n)$ for any n. In this case the space C_h and the corresponding quadratic form are very awkward to describe without the use of Fourier transforms.

2. Variational characterization and error estimate.

In this section we use the space C_h described by Theorem 1.1 to obtain a variational formulation of the interpolation problem we started with. We also look at certain bounds on the difference between a function $f \in C_h$ and an interpolant s that matches f on a set of points $x_1, \ldots, x_N$.

We begin with the question of whether C_h can interpolate arbitrary data.

Proposition 2.1: *Fix* $m \geq 0$ *and* h *in* $Q_m(\mathbb{R}^n)$. *Let* $\{x_1, \ldots, x_N\}$ *be a finite subset of* $\mathbb{R}^n$. *The following statements are equivalent*:

(a) *For every* v *in* $\mathbb{C}^N$ *there is an* f *in* C_h *that satisfies*

$$f(x_i) = v_i, \quad i=1,\ldots,N.$$

(b) $\sum_{i,j=1}^{N} \overline{c_i}\, c_j\, h(x_j - x_i) > 0$, *for every* $(c_1, \ldots, c_N) \neq 0$ *that satisfies* (1.1b),

(c) (1.1) *can be solved for every* v *in* $\mathbb{C}^N$.

Proof : To see that (a) implies (b), suppose (b) is not true. Then for some $c_1, \ldots, c_N$ as above we have $\sum_{i,j=1}^{N} \overline{c_i}\, c_j\, h(x_j - x_i) = 0$.

From (1.3) we see that $(\nu*h,\nu*h)_h = 0$ where $\nu = \sum_{i=1}^{N} c_i \delta_{x_i}$.

By part (c) of Theorem 1.1 this implies that $\nu(\bar{f}) = 0$ for all f in C_h. But then $\sum_{i=1}^{N} \overline{c_i} f(x_i) = 0$ and it is impossible to have $f(x_i) = c_i$. This is contrary to (a).

That (b) implies (c) was noted in the introduction (just before the definition of Q_m).

If c_i and k_α solve (1.1) then (1.1a) says $s(x_i) = v_i$. Here $s = \nu*h + p$ with $\nu \sum_{i=1}^{N} c_i \delta_{x_i}$ and $p(x) = \sum_{|\alpha|<m} k_\alpha x^\alpha$. By (1.1b), ν is in $P_{m-1}^{\perp}(\mathbb{R}^n)$, so s is in the space C_h. From this we see that (c) implies (a).

The next result shows that solutions of (1.1) correspond to interpolants from C_h that have minimum norm.

<u>Theorem 2.2</u>: <u>Fix</u> $m \geq 0$ <u>and</u> h <u>in</u> $Q_m(\mathbb{R}^n)$. <u>Given</u> $v = (v_1,\dots,v_N)$ <u>in</u> $\mathbb{C}^N$ <u>and distinct points</u> $x_1,\dots,x_N$ <u>in</u> $\mathbb{R}^n$, <u>let</u> $V = \{f \in C_h : f(x_i) = v_i, i=1,\dots,N\}$. <u>Then</u> V <u>is non-empty if and only if there are constants</u> c_j <u>and</u> k_α <u>that satisfy</u> (1.1). <u>In that case</u>

$$(2.1) \qquad s(x) = \sum_{j=1}^{N} c_j\, h(x-x_j) + \sum_{|\alpha|<m} k_\alpha\, x^\alpha$$

<u>belongs</u> <u>to</u> V <u>and</u> <u>for</u> <u>every</u> f <u>in</u> V

$$(2.2) \qquad \|f\|_h^2 = \|s\|_h^2 + \|f-s\|_h^2.$$

<u>Proof</u>: Let M be the subspace of $\mathbb{C}^N$ that is orthogonal to the subspace $P = \{(p(x_1),\ldots,p(x_N)) : p \in P_{m-1}(\mathbb{R}^n)\}$, so that $\mathbb{C}^N = M \oplus P$. For c in $\mathbb{C}^N$ let $\nu_c = \sum_{i=1}^{N} c_i\, \delta_{x_i}$. Note that $c \in M$ iff $\nu_c \in P_{m-1}^\perp$.

Suppose V is not empty. Choose f_o in V and let $\nu_o * h$ be the orthogonal projection of f_o onto the subspace $W = \{\nu_c * h : c \in M\}$. Then $(w, f_o - \nu_o * h)_h = 0$ for all w in W. Equivalently, for all c in M, $(\nu_c * h, g_o)_h = 0$ where $g_o = f_o - \nu_o * h$. By part (c) of Theorem 1.1, $\sum_{i=1}^{N} c_i\, \overline{g_o(x_i)} = 0$ for all c in M. Since $P = M^\perp$ there must be a polynomial p in P_{m-1} such that $g_o(x_i) = p(x_i)$ for $i=1,\ldots,N$. Thus, the function $s = \nu_o * h + p$ satisfies $s(x_i) = f_o(x_i) = v_i$, $i=1,\ldots,N$. From this it is seen that a solution of (1.1) is provided by the coefficients associated with ν_o and p.

Conversely, if (1.1) admits a solution then $V \neq 0$ since the

function given by (2.1) will be in V. To establish (2.2) we need only show that $(s,f-s)_h = 0$. As noted in the previous proof, $s = \nu * h + p$ with ν in $P_{m-1}^{\perp}$ and p in P_{m-1}. By Theorem 1.1,

$$(s,f)_h = (\nu * h, f)_h = \nu(\bar{f})$$

and

$$(s,s)_h = (\nu * h, s)_h = \nu(\bar{s}).$$

Now $\nu(\bar{f}) = \nu(\bar{s})$ because $f(x_i) = s(x_i)$, for $i=1,\ldots,N$. Hence $(s,f-s)_h = 0$.

For the remainder of this section, we fix $m \geq 1$ and look at bounds on the size of $|s(x_o) - f(x_o)|$ when s is the result of solving (1.1) with values coming from a function f in C_h (so that $v_i = f(x_i)$, $i=1,\ldots,N$). The bounds we obtain will depend on how x_o is situated in relation to the interpolation points x_i; roughly speaking x must be close to a certain number of the points x_i in order to get a good bound on $|s(x_o) - f(x_o)|$. More precisely, x_o must be close to a subset $Y = Y(x_o)$ of $\{x_1,\ldots,x_N\}$ that is unisolvent for P_{m-1}. This set Y will have $m' = \dim P_{m-1}$ points and for each y in Y there will be a unique ℓ_y in P_{m-1} with $\ell_y(y) = 1$ and $\ell_y(y') = 0$ for y' in $Y^{\sim}(y)$.

Some additional notation will be helpful. If K is a subset of $\mathbb{R}^n$, let $\operatorname{diam} K = \sup\{|x-x'| : x,x' \in K\}$. Also, let $\|f\|_K = \sup\{|f(x)| : x \in K\}$. Put $E_{m-1}(h,\varepsilon) = \inf\{\|h-p\|_{B(\varepsilon)} : p \in P_{m-1}\}$ where $B(\varepsilon)$ is the ball of radius ε centered at 0. Recalling

standard results on polynomial approximation, we note that $E_{m-1}(h,\varepsilon) = O(\varepsilon^m)$ when h is sufficiently smooth.

Theorem 2.3: *Fix* $m \geq 1$ *and* h *in* $Q_m(\mathbb{R}^n)$. *Given* f *in* C_h *and* $x_1,\dots,x_N$ *in* $\mathbb{R}^n$, *put* $V = \{g \in C_h : g(x_i) = f(x_i),\quad i=1,\dots,N\}$ *and let* s *be an element of* V *with minimal* C_h *norm. For each point* x_o *in* $\mathbb{R}^n$ *choose a* P_{m-1} *unisolvent set* $Y(x_o) \subset \{x_1,\dots,x_N\}$.

Then

$$|f(x_o) - s(x_o)| \leq \|f\|_h\,[1+\Lambda(x_o)]\,(E_{m-1}(h,d))^{1/2}$$

where $d = \operatorname{diam}(x_o) \cup Y(x_o)$ *and* $\Lambda(x) = \sum_{y\in Y(x_o)} |\ell_y(x)|$ *is the Lebesgue function for* P_{m-1} *interpolation at the points of* $Y(x_o)$.

Proof: Consider the measure $\nu_{x_o}^Y$ defined by

$$\nu_{x_o}^Y(g) = g(x_o) - \sum_{y\in Y} \ell_y(x_o)g(y). \tag{2.3}$$

For every $y \in Y$, $\nu_{x_o}^Y(\ell_y) = 0$. Thus $\nu_{x_o}^Y$ is in $P_{m-1}^{\perp}$. Also, $\nu_{x_o}^Y(\overline{f-s}) = \bar{f}(x_o) - \bar{s}(x_o)$, since $f-s$ vanishes on $\{x_1,\dots,x_N\} \supset Y$.

Hence, Cauchy's inequality and (2.2) give

(2.4)

$$|f(x_o)-s(x_o)| = |(\nu_{x_o}^Y * h,\ f-s)_h| \le \|\nu_{x_o}^Y * h\|_h \ \|f-s\|_h \le \|f\|_h \ \|\nu_{x_o}^Y * h\|_h .$$

To complete the proof, we need only establish

$$\|\nu_{x_o}^Y * h\|_h^2 \le [1 + \Lambda(x_o)]^2 \ E_{m-1}(h,d). \tag{2.5}$$

To show this, let $\pi : C(\mathbb{R}^n) \longrightarrow P_{m-1}$ be defined by $\pi g = \sum_{y \in Y} g(y) \ell_y$. From (2.3), we have

$$\nu_{x_o}^Y(g) = \delta_{x_o}[I-\pi]g$$

where $I\varphi = \varphi$ and $\delta_{x_o}\varphi = \varphi(x_o)$. For $t \in R^n$, define $\tau_t : C(\mathbb{R}^n) \longrightarrow C(\mathbb{R}^n)$ by $\tau_t g(x) = g(x-t)$. Note that

$$\nu_{x_o}^Y * h = \tau_{x_o} h - \sum_{y \in Y} \ell_y(x_o) \ \tau_y h.$$

Hence, by (c) in Theorem 1.1

$$\|\nu_{x_o}^Y * h\|_h^2 = \delta_{x_o}[I-\pi] \left[\tau_{x_o}\bar{h} - \sum_{y \in Y} \ell_y(x_o)\tau_y \bar{h}\right] \tag{2.6}$$

$$\le (1 + \sum_{y \in Y} |\ell_y(x_o)|) \max_{t \in \{x_o\} \cup Y} |\delta_{x_o}[I-\pi]\tau_t h| .$$

For p in P_{m-1},

$$[I-\pi]\tau_t h = [I-\pi]\tau_t(h-p).$$

Using $|\delta_{x_o}[I-\pi]g| \le [1+\Lambda(x_o)]\,\|g\|_{\{x_o\}\cup Y}$ with $g = \tau_t(h-p)$ we get

$$|\delta_{x_o}[I-\pi]\tau_t h| \le [1+\Lambda(x_o)]\,\|\tau_t(h-p)\|_{\{x_o\}\cup Y}$$

If K is any subset of $\mathbb{R}^n$ and $t \in K$, then

$$\|\tau_t g\|_K \le \|g\|_{B(\mathrm{diam}\ K)}.$$

Taking $K = \{x_o\}\cup Y$ and $g = h-p$, we see

$$|\delta_{x_o}[I-\pi]\tau_t h| \le [1+\Lambda(x_o)]\,\|h-p\|_{B(d)}$$

for every t in $\{x_o\}\cup Y$ and every p in P_{m-1}. On the right side, $\|h-p\|_{B(d)}$ can be replaced by $E_{m-1}(h,d)$; then from (2.6) we obtain (2.5).

3. Spaces with a reproducing property.

This section is devoted to a proof of Theorem 1.1. We start with uniqueness. The following theorem characterizes the functions f that belong to C_h and also their norm $\|f\|_h$. This characterization depends only on the choice of h (and m and n of course) since it is based solely on the semi-norm given by (3.2). Hence any space C_h that satisfies Theorem 1.1 is uniquely determined by (n, m and) h.

Theorem 3.1: *Fix $m \geq 0$ and h in $Q_m(\mathbb{R}^n)$. Let $C_h(\mathbb{R}^n)$ be any space of functions on $\mathbb{R}^n$ that satisfies the conditions of Theorem 1.1. Then a complex valued function f on $\mathbb{R}^n$ is in $C_h(\mathbb{R}^n)$ iff there is a constant C(f) such that*

$$(3.1) \qquad |\nu(\overline{f})| \leq C(f)\|\nu\|$$

for all ν in $P_{m-1}^{\perp}(\mathbb{R}^n)$. Here $\|\ \|$ denotes the semi-norm on $P_{m-1}^{\perp}$ defined by

$$(3.2) \qquad \|\nu\| = \left\{ \sum_{x \in \mathbb{R}^n} \sum_{y \in \mathbb{R}^n} \overline{\nu(\{x\})}\, \nu(\{y\})\, h(x-y) \right\}^{1/2}.$$

Furthermore, if f is in C_h and $C_*(f)$ is the infimum of all the constants C(f) for which (3.1) holds, then $C_*(f) = \|f\|_h$.

Proof: If f is in C_h then by part (c) of Theorem 1.1

$$|\nu(\overline{f})| = |(\nu * h, f)_h| \leq \|\nu * h\|_h \, \|f\|_h.$$

From (1.3) and (3.2) we have $\|\nu * h\|_h = \|\nu\|$. Thus (3.1) holds with $C(f) = \|f\|_h$. This proves $C_*(f) \le \|f\|_h$ and the "only if" half of the assertion about (3.1).

To prove the other half of the theorem fix an f for which a constant $C(f)$ exists and observe that (3.1) holds with $C(f) = C_*(f)$. Let $H = C_h/P_{m-1}$ and write $(\cdot,\cdot)_H$ for the inner product on H induced by $(\cdot,\cdot)_h$. Define $L : P_{m-1}^{\perp} \longrightarrow H$ by $L\nu = \nu * h + P_{m-1}$. Then $\|L\nu\|_H = \|\nu * h\|_h = \|\nu\|$. Using (3.1) we obtain

$$(3.3) \qquad |\varphi(\nu)| \le C_*(f)\ \|L\nu\|_H$$

where $\varphi(\nu) = \nu(\bar{f})$ defines φ as a linear function on $P_{m-1}^{\perp}$. From (3.3), if $L\nu_1 = L\nu_2$ then $\varphi(\nu_1) = \varphi(\nu_2)$. Thus there is a unique function Φ on $W = L(P_{m-1}^{\perp})$ such that $\varphi(\nu) = \Phi(L\nu)$ for all ν in $P_{m-1}^{\perp}$. Clearly Φ is linear and satisfies $|\Phi(w)| \le C_*(f)\|w\|_H$ for all w in W.

Since H is a Hilbert space, there is a v in H with $\|v\|_H \le C_*(f)$ such that $\Phi(w) = (w,v)_H$ for all w in W. Choose g in C_h so that $v = g + P_{m-1}$. Then

$$\nu(\bar{f}) = \Phi(L\nu) = (L\nu, v)_H = (\nu * h, g)_h = \nu(\bar{g}).$$

But if $\nu(\bar{f}) = \nu(\bar{g})$ for all ν in $P_{m-1}^{\perp}$ then, by Lemma 3.2 below, $\bar{f} - \bar{g} \in P_{m-1}$. In that case, $f = g + p$ with $\bar{p}$ (and hence p) in P_{m-1}. We conclude that f is in C_h and

$$\|f\|_h = \|g\|_h = \|v\|_H \le C_*(f).$$

This completes the second half of the proof.

Lemma 3.2: Let f be a complex valued function on $\mathbb{R}^n$. If $\nu(f) = 0$ for all ν in $P_{m-1}^{\perp}(\mathbb{R}^n)$ then $f \in P_{m-1}(\mathbb{R}^n)$.

Proof: Let $m' = \dim P_{m-1}(\mathbb{R}^n)$ and select points $x_1, \ldots, x_{m'}$ in $\mathbb{R}^n$ so that a basis $\{\ell_1, \ldots, \ell_{m'}\}$ exists for P_{m-1} with $\ell_i(x_j) = \delta_{ij}$. For each x in $\mathbb{R}^n$ define ν_x by

$$\nu_x(\varphi) = \varphi(x) - \sum_{i=1}^{m'} \ell_i(x)\varphi(x_i).$$

Note that ν_x is in $P_{m-1}^{\perp}$, since $\nu_x(\ell_i) = 0$ for $i=1,\ldots,m'$. By hypothesis, $\nu_x(f) = 0$ for all x. Thus

$$f = \sum_{i=1}^{n} f(x_i)\, \ell_i$$

which completes the proof.

Existence of the space described by Theorem 1.1 will be proved by using the following.

Theorem 3.3: Given $m \ge 0$ and h in $Q_m(\mathbb{R}^n)$ there is a Hilbert space H and a one to one linear transformation U of H into $C(\mathbb{R}^n)/P_{m-1}(\mathbb{R}^n)$ such that for every ν in $P_{m-1}^{\perp}$ there is a vector η_ν in H that satisfies

(i) $\qquad U\eta_\nu = \nu * h + P_{m-1}$ <u>and</u>

(ii) $\qquad (\eta_\nu, \eta)_H = \nu(\bar{f})$ <u>whenever</u> $U\eta = f + P_{m-1}$.

<u>Proof of Theorem 1.1</u>: Take $C_h = \bigcup \{U\eta : \eta \in H\}$ and define a semi-inner product on C_h by

$$(f,g)_h = \left(U^{-1}\left[f+P_{m-1}\right], U^{-1}\left[g+P_{m-1}\right]\right)_H.$$

Since U is one to one we see that $(f,f)_h = 0$ iff $f \in P_{m-1}$. Thus C_h/P_{m-1} is the inner product space that results by identifying f and g whenever $\|f-g\|_h = 0$. Evidently, U gives a unitary isomorphism betwen H and C_h/P_{m-1}. This verifies parts (a) and (b) of Theorem 1.1. Part (c) follows easily from (i) and (ii). Since uniqueness was established by Theorem 3.1, the proof is complete.

<u>Proof of Theorem 3.3</u>: If ν and λ are measures on $\mathbb{R}^n$, and have support consisting of a finite number of points, we take

(3.4)
$$\begin{aligned}(\nu,\lambda) &= \int\int h(x-y)\, d\nu(y)\, \overline{d\lambda(x)} \\ &= \sum_{x,y} \overline{\lambda(x)}\, \nu(y)\, h(x-y).\end{aligned}$$

The resulting function is linear (conjugate linear) in its first (second) argument. For $\nu \in P_{m-1}^{\perp}$ we have $(\nu,\nu) \geq 0$ because h is in Q_m. Also, the polarization identity gives $(\lambda,\nu) = \overline{(\nu,\lambda)}$

for ν, λ in $P_{m-1}^{\perp}$. This shows that $(\cdot,\cdot)$ is a semi - inner product on $P_{m-1}^{\perp}$. Thus $P_{m-1}^{\perp}/N$, where $N = \{\nu \in P_{m-1}^{\perp} : (\nu,\nu) = 0\}$, is an inner product space. Let H be its Hilbert space completion and let $\pi : P_{m-1}^{\perp} \longrightarrow H$ be the map given by $\pi\nu = \nu+N$. Note that $(\pi\nu,\pi\nu)_H = (\nu,\nu)$.

As in the proof of Lemma 3.2, we choose a P_{m-1} unisolvent set of points $x_1,\ldots,x_{m'}$ in $\mathbb{R}^n$ and define $\nu_x \in P_{m-1}^{\perp}$ by

$$\nu_x = \delta_x - \sum_{i=1}^{m'} \ell_i(x)\delta_{x_i}$$

We define $U : H \longrightarrow C(\mathbb{R}^n)/P_{m-1}$ by $U\eta = f_\eta + P_{m-1}$ where

$$f_\eta(x) = (\eta,\pi\nu_x)_H. \tag{3.5}$$

To verify that f_η is continuous we note $\|\pi\nu_t - \pi\nu_z\|_H = \|\nu_t - \nu_z\|$ and use

$$\lim_{t\to z} \|\nu_t - \nu_z\| = 0. \tag{3.6}$$

To see this, write $\nu_t - \nu_z = \delta_t - \delta_z + \sigma^{t,z}$ where

$$\sigma^{t,z} = \sum_{i=1}^{m'} [\ell_i(z) - \ell_i(t)]\,\delta_{x_i}.$$

Using sesqui-linearity and $(\delta_x,\delta_y) = h(y-x)$,

$$\|\nu_t - \nu_z\|^2 = 2h(0) - h(z-t) - h(t-z) + (\sigma^{t,z}, \sigma^{t,z})$$

$$+ (\delta_t - \delta_z, \sigma^{t,z}) + (\sigma^{t,z}, \delta_t - \delta_z).$$

Now $(\delta_t - \delta_z, \sigma^{t,z}) = \sum_{i=1}^{m'} [\ell_i(z) - \ell_i(t)]\,[h(x_i - t) - h(x_i - z)]$ which tends to 0 as $t \longrightarrow z$. After similar analysis of other terms, we obtain (3.6).

Next we show that

$$\nu(\bar{f}_\eta) = (\pi\nu, \eta)_H \tag{3.7}$$

for all $\nu \in P_{m-1}^{\perp}$. Writing $\nu = \sum_x \nu(x)\,\delta_x$ and using $0 = \nu(\ell_i) = \sum_x \nu(x)\,\ell_i(x)$ we find

$$\nu = \sum_x \nu(x)\,\nu_x. \tag{3.8}$$

From (3.5), $\nu(\bar{f}_\eta) = \sum_x \nu(x)\,\overline{f_\eta(x)} = \sum_x \nu(x)\,(\pi\nu_x, \eta)_H$. Applying (3.8) to this, we get (3.7).

The map U is clearly linear. To check that it is one to one, suppose $U\eta = P_{m-1}$. Then f_η is in P_{m-1} and from (3.7) we have $0 = (\pi\nu, \eta)_H$ for all ν in $P_{m-1}^{\perp}$. Since the image of π is dense in H, we conclude tht $\eta = 0$. Thus U is one to one.

To complete the proof we show that (i) and (ii) hold for

$\eta_\nu = \pi\nu$. From (3.5),

$$f_{\eta_\nu}(z) = (\pi\nu, \pi\nu_z)_H = (\nu, \nu_z)$$

$$= \sum_{x,y} \overline{\nu_z(x)}\ \nu(y)\ h(x-y)$$

$$= \sum_{x,y} \nu(y) \left[\delta_z(x) - \sum_{i=1}^{m'} \ell_i(z)\delta_{x_i}(x)\right] h(x-y)$$

$$= \sum_{y} \nu(y) \left[h(z-y) - \sum_{i=1}^{m'} \ell_i(z)h(x_i-y)\right].$$

Thus $f_{\eta_\nu} = \nu * h + p_\nu$ which gives (i). If $U\eta = f + P_{m-1}$ then $f = f_\eta + p$ and (ii) follows from (3.7).

References

[1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, The Theory of Splines and Their Applications, Academic Press, New York, 1967.

[2] N. Aronszajn, Theory of reproducing kernels, Tran. Amer. Math. Soc. 68 (1950), 337 - 404.

[3] J. Duchon, Interpolation des Fonctions de Deux Variables Suivant le Principe de la Flexion des Plaques Minces, RAIRO Analyse Numerique, 10 (1976), 5 - 12.

[4] J. Duchon, Splines Minimizing Rotation-Invariant Semi-norms in Sobolev Spaces, pp. 85 - 100 in Constructive Theory of Functions of Several Variables, Oberwolfach 1976, W. Schempp and K. Zeller, ed., Springer-Verlag, Berlin-Heidelberg, 1977.

[5] J. Duchon, Sur ℓ'Erreur d'Interpolation des Fonctions de Pusieurs Variables par les D^m-splines, RAIRO Analyse Numerique 12 (1978), 325 - 334.

[6] R. Franke, Scattered data interpolation: Tests of some methods, Mathematics of Computation 38 (1982) 181 - 199.

[7] R. Franke, Recent advances in the approximation of surfaces from scattered data, Naval Postgraduate School Technical Report, Monterey, California, 1986.

[8] I. M. Gelfand and N. Ya. Vilenkin, Generalized Functions, Vol.4, Academic Press, New York, 1964.

[9] R. L. Harder and R. N. Desmarais, Interpolation Using Surface Splines, J. Aircraft 9 (1972), 189 - 191.

[10] R. L. Hardy, Multiquadric equations of topography and other irregular surfaces, J. Geophys. Res. 76 (1971), 1905 - 1915.

[11] J. L. Kelly and I. Namioka, Linear Topological Spaces, D. Van Nostrand, Princeton, 1963.

[12] W. R. Madych and S. A. Nelson, Multivariate interpolation: a variational theory, manuscript, 1983.

[13] C. A. Micchelli, Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions, Constr. Approx 2(1986), 11 - 22.

[14] J. Meinguet, Multivariate Interpolation at Arbitrary Points Made Simple, J. of Applied Math. and Phys. (ZAMP) 30 (1979), 292 - 304.

[15] I. J. Schoenberg, Metric spaces and completely monotone functions, Ann. of Math. 39, 811 - 841.

[16] F. Utreras, Cross-Validation Techniques for Smoothing Spline Functions in One or Two Dimensions. pp. 196 - 232 in Smoothing Techniques for Curve Estimation, M. Rosenblatt and Th. Gasser, ed., Springer-Verlag, Berlin-Heidelberg, 1979.

Polyharmonic Cardinal Splines

W. R. Madych* S. A. Nelson†

Abstract

Polyharmonic splines, sometimes called thin plate splines, are distributions which are annihilated by iterates of the Laplacian in the complement of a a discrete set in Euclidean n-space and satisfy certain continuity conditions. The term cardinal is often used when the set of knots is a lattice. Here, in addition to developing certain basic properties of polyharmonic cardinal splines, it is shown that such splines interpolate numerical data on the lattice uniquely.

1 Introduction.

From one point of view, I. J. Schoenberg's theory of univariate cardinal splines of odd order, see [15], can be regarded as a development of certain properties of those functions, f, which satisfy

$$\frac{d^{2k} f}{dx^{2k}} = 0 \tag{1}$$

on the complement of the integer lattice Z and enjoy appropriate smoothness conditions on all of the real line R. A natural extension of these ideas to the multi-variate case would be to consider those functions f which satisfy differential equations analogous to (1) on the complement of the integer lattice Z^n in R^n and enjoy certain regularity properties on all of R^n. In this development we consider the case where the differential operators are powers of the Laplacian. Many of the results in [15] have appropriate analogues in this case. In this paper we consider only the basic properties of such splines together with the problems of existence and uniqueness of cardinal interpolation.

The motivation for our work came from an attempt to obtain a 'B spline like' basis for certain global interpolation schemes in R^n. More precisely, given a

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Both authors were partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

†Department of Mathematics, Iowa State University, Ames, IA 50011.

collection of points $x_1, \ldots, x_m$, in R^n, consider the linear subspace of functions, f, of the form

$$f(x) = \sum_{j=1}^{m} a_j \phi(x - x_j) \tag{2}$$

where ϕ is some fixed smooth function and $a_1, \ldots, a_m$ are arbitrary coefficients; call this subspace V_ϕ. Such functions are natural and simple candidates for interpolants of multivariate scattered data. In several interesting examples, such as $\phi(x) = |x|$ or $\phi(x) = \sqrt{1+|x|^2}$, the function ϕ does not decay at infinity; see [7,8] for these and other examples. As one may suspect, this causes various problems, both practical and theoretical. However, it is not difficult to see that, at least for the examples mentioned above, certain linear combinations of translates of ϕ, namely functions of form (2), decay at infinity rather quickly. It was hoped that such combinations would form a nice basis for V_ϕ analogous to that formed by the B splines in the classic univariate examples. In attempting to formulate a tractable theory we were led to consider the case where the set of points $x_1, \ldots, x_m$ becomes the integer lattice Z^n in R^n and the functions ϕ are fundamental solutions to certain powers of the Laplacian. This motivated us to examine Schoenberg's work, [14,15], more carefully and resulted in the development introduced here.

The idea of interpolating in terms of linear combinations of Green's functions to powers of the Laplacian is not new. Although the earliest published work devoted to the subject seems to be [10], it is quite clear that many mathematicians were aware of the idea and many of its consequences, either from the reproducing kernel Hilbert space viewpoint or from the transparent generalization of the variational aspect of univariate spline theory to the multivariate case; for instance see [9]. For examples of more recent work see [6,7,8,13]. Among the shortcomings of the obvious theory were the fact that such 'splines' do not have a localized basis and the restriction to the case of the finite domain. In [6] Duchon developed a variational theory for interpolation to all of R^n involving a finite number of constraints which overcomes the second mentioned shortcoming in an elegant way and allows for interesting generalizations. Although our development concerns an infinite number of constraints and does not rely on the variational properties of splines, it may also be regarded as a certain generalization of [6].

While working on an early draft of our theory we discovered that the subject of 'cardinal spline interpolation' was a booming business; for example see [3], [4], and the references cited there. (In addition to references on multivariate cardinal splines, the extensive bibliography in [3] contains many references to work involving the Green's function type splines mentioned above.) However, although some of our results may have analogues in the works cited above, we felt that this development had enough novelty to warrant completion.

This paper is organized as follows. In Section 2 we give the definitions and derive the basic properties of k-harmonic splines. The cardinal interpolation

problem for k-harmonic splines is taken up in Section 3 where it is shown that under certain circumstances this problem has a unique solution. The relationship of these splines to the multivariate analogue of the Whittaker cardinal series is indicated at the end of that section; this was prompted by a question raised by C. de Boor.

We use mathematical notation which is standard when dealing with multivariate functions. For example, the symbols μ and ν usually will denote multi-indexes, $x^\mu = x_1^{\mu_1} \dots x_n^{\mu_n}$, etc.; $\mathcal{S}(R^n)$ and $\mathcal{S}'(R^n)$ denote the Schwartz space of rapidly decreasing functions and its dual, the space of tempered distributions. The introductory chapter of [11] contains a concise summary of this so-called multi-index notation and basic facts on distributions and Fourier transforms; other references which contain basic material used here concerning distributions, Fourier transforms, and several complex variables are [2], [5], and [16]. For the Fourier transform we use a standard normalization which is slightly different from that used in [11], namely,

$$\hat{\phi}(\xi) = (2\pi)^{-n/2} \int e^{-i(\xi,x)} \phi(x)dx$$

when ϕ is in $\mathcal{S}(R^n)$; here the integral is taken over all of R^n and $\hat{\phi}$ is the Fourier transform of ϕ. In what follows, integrals, as in the above case, are taken over R^n unless specifically denoted otherwise.

2 Definitions and basic properties.

Recall that a function or distribution u is said to be *k-harmonic*, k a positive integer, if

$$\Delta^k u = 0 \tag{3}$$

on R^n. Here Δ is the usual Laplace operator defined by

$$\Delta u = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2}$$

and, if k is greater than one, Δ^k denotes its k-th iterate, $\Delta^k u = \Delta(\Delta^{k-1}u)$. Of course $\Delta^1 = \Delta$. A *polyharmonic* function is one which is k-harmonic for some positive integer k; for a treatise on the subject see [1].

The class $H_k(R^n)$ is the collection of all k-harmonic tempered distributions. In what follows the classes $H_k(R^n)$ play a role similar to that played by the polynomials, $\Pi_{2k-1}(R)$, in univariate spline theory. Indeed we have the following.

Proposition 1 *The class $H_k(R^n)$ is a subspace of polynomials which contains $\Pi_{2k-1}(R^n)$.*

Proof If u satisfies (3) then its Fourier transform, $\hat{u}$, satisfies $|\xi|^{2k}\hat{u}(\xi) = 0$ for all ξ in R^n. It follows that $\hat{u}$ is a distribution supported at the origin and thus must be a finite linear combination of the Dirac distribution and its derivatives. Hence u must be a polynomial. That $H_k(R^n)$ contains $\Pi_{2k-1}(R^n)$ follows from the fact that Δ^k is a homogeneous differential operator of order $2k$. ■

For k a positive integer satisfying $2k \geq n+1$, we define $SH_k(R^n)$ to be the subspace of $S'(R^n)$ whose elements f enjoy the following properties:

$$(4) \qquad \begin{array}{ll} (i) & f \text{ is in } C^{2k-n-1}(R^n) \text{ and} \\ (ii) & \Delta^k f = 0 \text{ on } R^n \setminus Z^n. \end{array}$$

Here Z denotes the set of integers and Z^n denotes the integer lattice in R^n. Elements of Z^n are denoted by boldface symbols such as $\mathbf{j}$ and $\mathbf{m}$.

A *k-harmonic cardinal spline* is an element of $SH_k(R^n)$. We say that a function or distribution is a *polyharmonic cardinal spline* if it is in one of the classes $SH_k(R^n)$.

The reason for the condition $2k \geq n+1$ can be explained as follows. If $2k < n+1$ any distribution which is locally in L^∞ and satisfies condition (4ii) must be k-harmonic on R^n. Thus if $SH_k(R^n)$ is to consist of functions which satisfy (4ii) and for which pointwise evaluation on Z^n makes sense then the last observation implies that $SH_k(R^n)$ must in fact be the space $H_k(R^n)$. Like the space of polynomials in the case $n=1$, the class $H_k(R^n)$ can only interpolate very restricted data on Z^n. Since we wish the space $SH_k(R^n)$ to be rich enough to interpolate a wide class of data on Z^n, we assume $2k \geq n+1$.

In what follows we always assume that $2k \geq n+1$. For the sake of clarity however, we will remind the reader of this from time to time.

Fundamental solutions of (3) play an important role in the description of $SH_k(R^n)$. For future reference we define E_k to be the fundamental solution of (3) given by

$$(5) \qquad E_k(x) = \begin{cases} c(n,k)|x|^{2k-n} & \text{if } n \text{ is odd} \\ c(n,k)|x|^{2k-n}log|x| & \text{if } n \text{ is even} \end{cases}$$

where $c(n,k)$ is a constant which depends only on n and k and is chosen so that $\Delta^k E_k(x) = \delta(x)$. Here $\delta(x)$ denotes the unit Dirac distribution at the origin. The Fourier transform of E_k is

$$(6) \qquad \hat{E}_k(\xi) = (2\pi)^{-n/2}(-|\xi|^2)^{-k}.$$

The following propositions give some of the basic properties of the spaces $SH_k(R^n)$.

Proposition 2 *If f is a tempered distribution then the following conditions are equivalent:*

(i) f is in $SH_k(R^n)$.

(ii) f satisfies

(7) $$\Delta^k f(x) = \sum a_{\mathbf{j}} \delta(x - \mathbf{j})$$

where the a's are constants and the sum is taken over all $\mathbf{j}$ *in* Z^n.

Proof Suppose f is in $SH_k(R^n)$. To see that f satisfies (7) observe that in a sufficiently small neighborhood N of any point $\mathbf{j}$ we have

$$\Delta^k f(x) = \sum_{\nu} b_\nu D^\nu \delta(x - \mathbf{j})$$

where the sum is taken over some finite set of multi-indicies ν. Representation (7) will follow if we can show that $b_\nu = 0$ for $\nu \neq 0$. To see this let ϕ be any infinitely differentiable function with support in N such that $\phi(x) = 1$ in a neighborhood of $\mathbf{j}$. Then, since f is in $C^\infty(R^n \setminus Z^n)$,

$$\Delta^k(\phi f) = \sum_{\nu} b_\nu D^\nu \delta(x - \mathbf{j}) + \psi$$

where ψ is an infinitely differentiable function with support in N. Now, if E is any fundamental solution of (3) then

$$\phi f = E * (\Delta^k \phi f) = \sum_{\nu} b_\nu D^\nu E(x - \mathbf{j}) + E * \psi$$

where $E * \psi$ is infinitely differentiable on R^n. The fact that $b_\nu = 0$ for $\nu \neq 0$ follows from the last equation, the fact that ϕf is in $C^{2k-n-1}(R^n)$, and the behavior of $D^\nu E(x - \mathbf{j})$ in a neighborhood of $\mathbf{j}$.

To see that (7) implies that f is in $C^{2k-n-1}(R^n)$ observe that in any sufficiently small neighborhood N of $\mathbf{j}$ $f(x) - a_{\mathbf{j}} E(x - \mathbf{j})$ is equal to a k-harmonic function and therefore $f(x) - a_{\mathbf{j}} E(x - \mathbf{j})$ is in $C^\infty(N)$. Note that $E(x - \mathbf{j})$ is in $C^{2k-n-1}(R^n)$ and hence it follows that f is in $C^{2k-n-1}(N)$. The desired conclusion is now an easy consequence of this fact. ∎

In what follows we say that a sequence $a_{\mathbf{j}}$, $\mathbf{j}$ in Z^n, is of polynomial growth if there are constants c and p such that $|a_{\mathbf{j}}| \leq c(1+|\mathbf{j}|)^p$ for all $\mathbf{j}$. Similarly a locally bounded function f is said to be of polynomial growth if $|f(x)| \leq c(1+|x|)^p$ for all x in R^n. The following observation is an easy consequence of the definitions.

Proposition 3 *If f is in $SH_k(R^n)$ then the coefficients in representation (7) are unique and are of polynomial growth. Furthermore, if f_1 and f_2 are in $SH_k(R^n)$ and $\Delta^k f_1 = \Delta^k f_2$ then $f_1 - f_2$ is in $H_k(R^n)$.*

The proof of the following proposition is somewhat technical and perhaps disruptive to the flow of main ideas of this section. We include it for completeness.

Proposition 4 *If f is in $SH_k(R^n)$ then f is of polynomial growth.*

Proof For ϕ in $\mathcal{S}(R^n)$ let

$$|\phi|_{\mu,\nu} = \sup|x^\mu D^\nu \phi(x)|$$

where the supremum is taken over all x in R^n. Since f is in $\mathcal{S}'(R^n)$ there are constants M, N, and C_1 so that

$$|\langle f, \phi\rangle| \leq C_1 \sum |\phi|_{\mu,\nu} \tag{8}$$

where the sum is taken over all multi-indexes μ and ν such that $0 \leq |\mu| \leq M$ and $0 \leq |\nu| \leq N$. Since f is in $SH_k(R^n)$ there are constants M_1 and C_2 such that

$$|\langle \Delta^k f, \phi\rangle| \leq C_2 \sum |\phi|_{\mu,0} \tag{9}$$

where the sum is taken over all multi-indexes μ such that $0 \leq |\mu| \leq M_1$.

Observe that $\langle f, \phi\rangle = \langle f, \phi_1\rangle + \langle (-\Delta)^k f, \phi_1\rangle$ and by induction

$$\langle f, \phi\rangle = \langle f, \phi_m\rangle + \sum_{j=1}^{m} \langle (-\Delta)^k f, \phi_j\rangle \tag{10}$$

where ϕ_j is defined by the formula for its Fourier transform

$$\hat{\phi}_j(\xi) = (1 + |\xi|^{2k})^{-j} \hat{\phi}(\xi) \tag{11}$$

Formula (11) together with properties of the Fourier transform imply that

$$x^\mu \phi_j(x) = \sum_{\nu \leq \mu} g_\nu * \psi_\nu(x) \tag{12}$$

where ψ_ν is defined by $\psi_\nu(x) = x^\nu \phi(x)$ and the g_ν are bounded continuous functions. Hence there is a constant C_3 such that

$$|\phi_j|_{\mu,0} \leq C_3 \sum_{\nu \leq \mu} \int |x^\nu \phi(x)| dx. \tag{13}$$

Similar reasoning shows that if m satisfies $2k(m-1) > |\nu|$ then

$$|\phi_m|_{\mu,\nu} \leq C_4 \sum_{\kappa \leq \mu} \int |x^\kappa \phi(x)| dx \tag{14}$$

where C_4 is a constant independent of ϕ.

Formula (10) together with estimates (8), (9), (13), and (14) imply that there are positive constants C and α, independent of ϕ, such that

$$|\langle f,\phi\rangle| \leq C\int(1+|x|)^{\alpha}|\phi(x)|dx \tag{15}$$

for all ϕ in $\mathcal{S}(R^n)$. Estimate (15) together with Riesz representation imply the desired result. ■

This last proposition, when combined with the fact that continuous functions of polynomial growth are distributions in $\mathcal{S}'(R^n)$ shows that $SH_k(R^n)$ could be defined, without any loss of generality, as the class of continuous functions which satisfy (4) and are of polynomial growth.

In the case $n = 1$ $SH_k(R^n)$ is the subspace of S_m, $m = 2k-1$, consisting of functions of polynomial growth. The space S_m is the space of piecewise polynomial functions defined by Schoenberg, see [15]. The reason we restrict our attention to $\mathcal{S}'(R^n)$ is because our development relies on the use of the Fourier transform. Thus in spirit this development is similar to that in Schoenberg's earlier work [14].

If α is a real parameter we define $SH_k^{\alpha}(R^n)$ to be that subspace of $SH_k(R^n)$ consisting of functions f which satisfy $|f(x)| \leq c(1+|x|)^{\alpha}$ for some constant c. As an immediate consequence of Proposition 4 we have the following.

Corollary 1 $SH_k(R^n) = \bigcup SH_k^{\alpha}(R^n)$ *where the union is taken over all real* α.

To complete our list of basic properties of $SH_k(R^n)$ we include the following.

Proposition 5 *Given a sequence* $\{a_{\mathrm{j}}\}$, j *in* Z^n, *which is of polynomial growth, there is an* f *in* $SH_k(R^n)$ *such that*

$$\Delta^k f(x) = \sum a_{\mathrm{j}}\delta(x-\mathrm{j}). \tag{16}$$

Proof Recall that $\Delta^k u = v$ has a solution u in $\mathcal{S}'(R^n)$ whenever v is in $\mathcal{S}'(R^n)$; for example see [12]. The hypothesis implies that the right hand side of (16) is a tempered distribution and thus there is a tempered distribution f such that (16) holds. That f is in $SH_k(R^n)$ now follows from Proposition 2. ■

Proposition 6 $SH_k(R^n)$ *is a closed subspace of* $\mathcal{S}'(R^n)$.

Proof The mapping $f \rightarrow \Delta^k f$ is continuous as an operator from $\mathcal{S}'(R^n)$ to $\mathcal{S}'(R^n)$, which means that the inverse image of a closed set is closed under this mapping. Now, the set of those elements in $\mathcal{S}'(R^n)$ consisting of Radon measures supported on Z^n is a closed subspace of $\mathcal{S}'(R^n)$, and, since $SH_k(R^n)$ is the inverse image of this subspace under the mapping described above, the desired result follows. ■

3 Cardinal spline interpolation

The problem of *cardinal interpolation for k-harmonic splines* is the following:

> Given a sequence of real or complex numbers, $v = \{v_{\mathrm{j}}\}$, j in Z^n, find an element f in $SH_k(R^n)$ such that $f(\mathrm{j}) = v_{\mathrm{j}}$ for all j.

This, of course, is a simple generalization of the standard problem in the univariate case, see [15, p. 33].

Since elements of $SH_k(R^n)$ are of polynomial growth it is clear that in order for this problem to have a solution a necessary requirement on the sequence v is that it also be of polynomial growth. This requirement is also sufficient, as we shall show. We begin by first considering the fundamental functions of interpolation.

The function L_k is defined by the formula for its Fourier transform:

$$\hat{L}_k(\xi) = (2\pi)^{-n/2} \frac{|\xi|^{-2k}}{\sum_{\mathrm{j}\in Z^n} |\xi - 2\pi \mathrm{j}|^{-2k}}. \tag{17}$$

If k is an integer such that $2k \geq n+1$ then $L_k(\xi)$ is well defined as an absolutely convergent integral. This function is the ***fundamental function of interpolation*** for k-harmonic splines. The reason for this terminology is the fact that $L_k(\mathrm{j}) = \delta_{0\mathrm{j}}$, where $\delta_{0\mathrm{j}}$ is the Kronecker delta. This and other properties of L_k and some of the consequences for the interpolation problem are stated in Propositions 7 and 8. Before turning to these propositions however, we need to consider certain technical properties and lemmas concerning L_k and the related periodic distribution, $\hat{\Phi}_k$, defined by

$$\hat{\Phi}_k(\xi) = (-|\xi|^2)^k \hat{L}_k(\xi). \tag{18}$$

Given a subset $\mathcal{A}$ of the real line R and a positive number ϵ, $\mathcal{A}_\epsilon$ is the subset of the complex plane defined by

$$\mathcal{A}_\epsilon = \{\tau \in C \,:\, \Re\tau \in \mathcal{A} \text{ and } -\epsilon < \Im\tau < \epsilon\}$$

where $\Re\tau$ and $\Im\tau$ are the real and imaginary parts of τ respectively. Similarly $\mathcal{A}_\epsilon^n = \mathcal{A}_\epsilon \times \cdots \times \mathcal{A}_\epsilon$ is a subset of C^n. The symbol Q denotes the interval $-\pi < \rho \leq \pi$ and Q^n denotes the cube

$$Q^n = \{\xi = (\xi_1, \ldots, \xi_n) \,:\, -\pi < \xi_j \leq \pi,\ j = 1, \ldots, n\}.$$

Lemma 1 *The functions $\hat{\Phi}_k$ and $\hat{L}_k$ have extensions which are analytic in a tube R_ϵ^n, for some $\epsilon > 0$.*

Proof For ξ and η in R^n, let

$$\zeta = (\zeta_1, \ldots, \zeta_n) = (\xi_1 + i\eta_1, \ldots \xi_n + i\eta_n) = \xi + i\eta$$

denote a point in complex n-space, C^n, and put

$$q(\zeta) = -\sum_{m=1}^{n} \zeta_m^2.$$

Observe that

$$(2\pi)^{n/2}[\hat{\Phi}_k(\xi)]^{-1} = \{1 + [q(\xi)]^k F(\xi)\}[q(\xi)]^{-k} \tag{19}$$

where

$$F(\zeta) = \sum_{\mathbf{j} \in Z^n \setminus \{0\}} q(\zeta - 2\pi \mathbf{j})^{-k}$$

Choose ϵ small enough so that $q(\zeta) \neq 0$ for all ζ in $R_\epsilon^n \setminus Q_\epsilon^n$. Then it is readily checked that F is analytic in Q_ϵ^n. For ξ in Q^n, $[q(\xi)]^k F(\xi) \geq 0$. Thus, by reducing ϵ if necessary, we may assume that $1 + [q(\zeta)]^k F(\zeta)$ has no zeros in Q_ϵ^n. Now, from (19), it follows that $\hat{\Phi}_k$ extends analytically to Q_ϵ^n and hence, by periodicity, it extends analytically to R_ϵ^n.

The analytic extension of $\hat{L}_k$ is given by $\hat{\Phi}_k(\zeta)[q(\zeta)]^{-k}$. From (19) it is evident that this extension is analytic in Q_ϵ^n. Analyticity on the rest of R_ϵ^n is clear from the fact that $q(\zeta) \neq 0$ there. ■

Lemma 2 *The Fourier transform of $\hat{\Phi}_k$ is a sum of constant multiples of translates of the delta function. More precisely, $\hat{\hat{\Phi}}_k = \Phi_k$ and*

$$\Phi_k(x) = \sum_{\mathbf{j} \in Z^n} a_{\mathbf{j}} \delta(x - \mathbf{j}) \tag{20}$$

where the $a_{\mathbf{j}}$'s depend on k and n. Furthermore, there are positive constants, C and c, which are independent of $\mathbf{j}$ such that

$$|a_{\mathbf{j}}| \leq C \exp(-c|\mathbf{j}|) \tag{21}$$

for all $\mathbf{j}$.

Proof The periodicity of $\hat{\Phi}_k$ implies that Φ_k satisfies (20) with

$$a_{\mathbf{j}} = (2\pi)^{-n/2} \int_{Q^n} e^{i\langle \mathbf{j}, \xi \rangle} \hat{\Phi}_k(\xi) d\xi. \tag{22}$$

Now, by virtue of analyticity of $\hat{\Phi}_k$ the set Q^n in (22) can be replaced by $\{\zeta : \Re\zeta \in Q^n \text{ and } \Im\zeta = \gamma\}$ where $\gamma = (\gamma_1, \ldots, \gamma_n)$, γ_m are constants, $|\gamma_m| = \epsilon/2$, $m = 1, \ldots, n$, and the sign of γ_m is chosen so that $\langle \mathbf{j}, \gamma \rangle > 0$ whenever $\mathbf{j} \neq 0$; here ϵ is the same as that in Lemma 1. Upon doing this, a routine estimate of the resulting integral gives (21). ■

Proposition 7 *Let L_k be defined by the formula for its Fourier transform (17), where k is an integer which satisfies $2k \geq n+1$. Then L_k has the following properties:*

(i) L_k is a k-harmonic cardinal spline.

(ii) For all $\mathbf{j}$ in Z^n

$$L_k(\mathbf{j}) = \begin{cases} 1 & \text{if } \mathbf{j} = 0 \\ 0 & \text{if } \mathbf{j} \neq 0 \end{cases} \tag{23}$$

(iii) There are positive constants A and a, depending on n and k but independent of x, such that

$$|L_k(x)| \leq A \exp(-a|x|) \tag{24}$$

for all x in R^n.

(iv) L_k has the following representations in terms of E_k:

$$L_k(x) = \sum_{\mathbf{j} \in Z^n} a_{\mathbf{j}} E_k(x - \mathbf{j}) = \Phi_k * E_k(x) \tag{25}$$

where Φ_k is the function whose Fourier transform is defined by (18) and the $a_{\mathbf{j}}$'s are the constants defined in Lemma 2. The series converges absolutely and uniformly on all compact subsets of R^n.

Proof *(i)* This follows from (20) and the fact that $\Delta^k L_k = \Phi_k$.

(ii) Observe that $\sum_{\mathbf{j}} \hat{L}_k(\xi - 2\pi\mathbf{j}) = (2\pi)^{-n/2}$, and write

$$L_k(\mathbf{j}) = (2\pi)^{-n/2} \int \hat{L}_k(\xi) e^{i(\mathbf{j},\xi)} d\xi = (2\pi)^{-n/2} \sum_{\mathbf{m} \in Z^n} \int_{Q^n} \hat{L}_k(\xi - 2\pi\mathbf{m}) e^{i(\mathbf{j},\xi)} d\xi$$

$$= (2\pi)^{-n/2} \int_{Q^n} \sum_{\mathbf{m} \in Z^n} \hat{L}_k(\xi - 2\pi\mathbf{m}) e^{i(\mathbf{j},\xi)} d\xi = \frac{1}{(2\pi)^n} \int_{Q^n} e^{i(\mathbf{j},\xi)} d\xi.$$

The desired result is now an immediate consequence of the last formula.

(iii) The proof of this estimate is analogous to the proof of (21) in Lemma 2.

(iv) This is a transparent consequence of the definitions and Lemma 2. ■

For later reference, we define $\mathcal{Y}^\alpha$, α real, to be the collection of those sequences $v = \{v_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, which satisfy

$$|v_{\mathbf{j}}| \leq c(1 + |\mathbf{j}|)^\alpha \tag{26}$$

for some constant c. Note that these classes are analogous to the ones in [15, p. 34]; they will also be used here in a similar fashion.

Proposition 8 *Suppose $v = \{v_{\mathrm{j}}\}$, j in Z^n, is a sequence of polynomial growth and f_v is the function defined by*

$$f_v(x) = \sum_{\mathrm{j}\in Z^n} v_{\mathrm{j}} L_k(x-\mathrm{j}). \tag{27}$$

Then the following is true:

(i) The expansion (27) converges absolutely and uniformly in every compact subset of R^n.

(ii) The function f_v is a k-harmonic spline and $f_v(\mathrm{j}) = v_{\mathrm{j}}$ for all j.

(iii) If v is in $\mathcal{Y}^\alpha$ then f_v is in $SH_k^\alpha(R^n)$.

(iv) If $v_{\mathrm{j}} = P(\mathrm{j})$ for all j, *where $P(x)$ is a k-harmonic polynomial then $f_v(x) = P(x)$ for all x in R^n.*

Proof Items *(i)-(iii)* are immediate consequences of Proposition 7.
(iv) Suppose P is a k-harmonic polynomial and f is defined by

$$f(x) = \sum_{\mathrm{j}\in Z^n} P(\mathrm{j}) L_k(x-\mathrm{j}).$$

Recall that any locally integrable function g of polynomial growth is in $\mathcal{S}'(R^n)$; furthermore for any ϕ in $\mathcal{S}(R^n)$ we have

$$\langle g,\phi\rangle = \int g(x)\phi(x)dx \text{ and } g*\phi(x) = \int g(y)\phi(x-y)dy.$$

To prove *(iv)* of this proposition it suffices to show that

$$\langle f,\phi\rangle = \langle P,\phi\rangle \tag{28}$$

for all ϕ in $\mathcal{S}(R^n)$.
To see (28) observe that

$$\langle f,\phi\rangle = \sum_{\mathrm{j}\in Z^n} P(\mathrm{j}) L_k*\phi(\mathrm{j}) = (2\pi)^{n/2}\sum_{\mathrm{j}\in Z^n}[P(iD)(\hat{L}_k(\xi)\hat{\phi}(\xi))]_{\xi=2\pi\mathrm{j}} \tag{29}$$

where the last equality follows by virtue of the Poisson summation formula. To evaluate the expression on the extreme right in (29) set

$$A_{\mathrm{j}} = [P(iD)(\hat{L}_k(\xi)\hat{\phi}(\xi))]_{\xi=2\pi\mathrm{j}}$$

and observe that for each j in Z^n $\hat{L}_k$ may be expressed as

$$(30) \qquad (2\pi)^{n/2}\hat{L}_k(\xi) = \begin{cases} 1+|\xi|^{2k}\psi_0(\xi) & \text{if } \mathbf{j}=0 \\ |\xi-2\pi\mathbf{j}|^{2k}\psi_{\mathbf{j}}(\xi) & \text{otherwise} \end{cases}$$

where the $\psi_{\mathbf{j}}$'s are smooth functions such that $\hat{\phi}_{\mathbf{j}}(\xi) = \psi_{\mathbf{j}}(\xi)\hat{\phi}(\xi)$ is a test function in $\mathcal{S}(R^n)$ for each j. In view of (30) we may write

$$(31) \qquad (2\pi)^{n/2}A_{\mathbf{j}} = \begin{cases} \int P(x)[\phi(x)+(-\Delta)^k\phi_0(x)]dx & \text{if } \mathbf{j}=0 \\ \int P(x)(-\Delta)^k[e^{-i(\mathbf{j},x)}\phi_{\mathbf{j}}(x)]dx & \text{otherwise.} \end{cases}$$

Now, integrating by parts and using the fact that $\Delta^k P = 0$ results in

$$(32) \quad \int P(x)(-\Delta)^k[e^{-i(\mathbf{j},x)}\phi_{\mathbf{j}}(x)]dx = \int[(-\Delta)^k P(x)]e^{-i(\mathbf{j},x)}\phi_{\mathbf{j}}(x)dx = 0$$

when $\mathbf{j} \neq 0$. Similar reasoning shows that

$$(33) \qquad (2\pi)^{n/2}A_0 = \int P(x)\phi(x)dx.$$

Formulas (29), (30), (31), (32) and (33) imply (28) which is the desired result.
■

It should now be clear why L_k is called the fundamental function of interpolation for k-harmonic splines. Note that items *(ii)* and *(iv)* of the previous proposition may be restated as follows.

Corollary 2 *If the data v is of polynomial growth then the cardinal interpolation problem for k-harmonic splines has a solution which has a unique representation in terms of fundamental functions of interpolation. The solution is given by*

$$(34) \qquad f_v(x) = \sum_{\mathbf{j}\in Z^n} v_{\mathbf{j}} L_k(x-\mathbf{j}).$$

We now take up the question of uniqueness by first considering the following technical lemma.

Lemma 3 *If f is in $SH_k(R^n)$ and f is Z^nperiodic then f is a constant.*

Proof The hypothesis implies that

$$\hat{f}(\xi) = \sum c_{\mathbf{j}}\delta(\xi-2\pi\mathbf{j}) \text{ and } \Delta^k f(x) = a\sum\delta(x-\mathbf{j}),$$

where both sums are taken over all $\mathbf{j}$ in Z^n; a and the $c_{\mathbf{j}}$'s are constants. These formulas together with Poisson summation imply

$$(-|2\pi \mathbf{j}|^2)^k c_{\mathbf{j}} = a \tag{35}$$

for all $\mathbf{j}$ in Z^n. The equation corresponding to $\mathbf{j} = 0$ in (35) implies that a is 0; the rest of the equations in system (35) imply that $c_{\mathbf{j}} = 0$ whenever $\mathbf{j} \neq 0$. Thus $f(x) = c_0$, which is the desired result. ■

Proposition 9 *If f is in $SH_k(R^n)$ and $f(\mathbf{j}) = 0$ for all $\mathbf{j}$ in Z^n then f is identically* 0.

Proof Recall that

$$\Delta^k f(x) = \sum_{\mathbf{j} \in Z^n} c_{\mathbf{j}} \delta(x - \mathbf{j}). \tag{36}$$

Now, if $c_{\mathbf{j}} = 0$ for all $\mathbf{j}$, f must be a k-harmonic polynomial. This together with the fact that $f(\mathbf{j}) = 0$ for all $\mathbf{j}$ implies the desired result.

Thus to complete the proof it suffices to show that $c_{\mathbf{j}} = 0$ for all $\mathbf{j}$.

To see this, write

$$g(x) = \sum_{\mathbf{j} \in Z^n} c_{\mathbf{j}} L_k(x - \mathbf{j}),$$

and observe that

$$\Delta^k(g - \Phi_k * f) = 0 \tag{37}$$

where Φ_k is the distribution defined by Lemma 2. From (37) it follows that $g - \Phi_k * f$ is a k-harmonic polynomial P. Since $\Phi_k * f(\mathbf{j}) = 0$ for all $\mathbf{j}$ it follows that $g(\mathbf{j}) = P(\mathbf{j}) = c_{\mathbf{j}}$.

Now, if $P \neq 0$, there is a finite difference operator, T, such that

$$TP(x) = \sum_{\mathbf{j} \in \mathcal{F}} b_{\mathbf{j}} P(x - \mathbf{j}) = B \tag{38}$$

where the sum is taken over a finite subset $\mathcal{F}$ of Z^n and B is a non-zero constant. Write

$$\Delta^k Tf(x) = T\Delta^k f(x) = \sum_{\mathbf{j} \in Z^n} TP(\mathbf{j})\delta(x - \mathbf{j}) = B \sum_{\mathbf{j} \in Z^n} \delta(x - \mathbf{j}) \tag{39}$$

and observe that (39) means that $Tf(x - \mathbf{j}) - Tf(x)$ is a k-harmonic polynomial for each $\mathbf{j}$. This, together with the fact that $Tf(\mathbf{m} - \mathbf{j}) - Tf(\mathbf{m}) = 0$ for all $\mathbf{m}$ and $\mathbf{j}$ implies that Tf is Z^n periodic. Now, by virtue of Lemma 3 Tf is a constant and hence

$$\Delta^k Tf = 0. \tag{40}$$

Formulas (39) and (40) contradict the fact that $B \neq 0$ and thus imply the desired result. ∎

An immediate consequence of Proposition 9 of course is the fact that any solution for the problem of cardinal interpolation for k-harmonic splines is unique. We summarize our results concerning the interpolation problem as follows:

Theorem 1 *If $v = \{v_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, is a sequence of polynomial growth then there is a unique k-harmonic spline f such that $f(\mathbf{j}) = v_{\mathbf{j}}$ for all $\mathbf{j}$. Furthermore, if v is in $\mathcal{Y}^\alpha$ then f is in $SH_k^\alpha(R^n)$.*

Note that the converse of the theorem above follows immediately from the definitions and results in Section 2.

Theorem 2 *Every k-harmonic spline f has a unique representation in terms of translates of L_k, namely*

$$f(x) = \sum_{\mathbf{j} \in Z^n} f(\mathbf{j}) L_k(x - \mathbf{j}).$$

We conclude this paper with a result that may justify the use of the adjective 'cardinal' when referring to these splines. For the univariate case see [15].

Proposition 10 *For $x = (x_1, \ldots, x_n)$*

$$\lim_{k \to \infty} L_k(x) = \prod_{j=1}^{n} \frac{\sin \pi x_j}{\pi x_j} \tag{41}$$

uniformly in x on R^n.

Proof First observe that the formula for $\hat{L}_k$ may be rewritten as

$$\hat{L}_k(\xi) = (2\pi)^{-n/2} \left\{ 1 + \sum_{\mathbf{j} \neq 0} \frac{|\xi|^{2k}}{|\xi - 2\pi \mathbf{j}|^{2k}} \right\}^{-1}. \tag{42}$$

Now, for $\mathbf{j} \neq 0$ and ξ in the interior of Q^n, $|\xi - 2\pi\mathbf{j}| > |\xi|$. Hence, for such ξ, it is clear from (42) that

$$\lim_{k \to \infty} \hat{L}_k(\xi) = (2\pi)^{-n/2}. \tag{43}$$

To see what happens for general ξ's write, by virtue of the periodicity of $\hat{\Phi}_k$,

$$\hat{L}_k(\xi) = (-|\xi|^2)^{-k} \hat{\Phi}_k(\xi - 2\pi\mathbf{j}) = \left\{ \frac{|\xi - 2\pi\mathbf{j}|}{|\xi|} \right\}^{2k} \hat{L}_k(\xi - 2\pi\mathbf{j}). \tag{44}$$

From (44) it is clear that for ξ in the complement of Q^n

$$\lim_{k\to\infty} \hat{L}_k(\xi) = 0. \tag{45}$$

Again using (44) together with routine estimates shows that, whenever $2k \geq n+1$, $\hat{L}_k$ is dominated by an integrable function, independent of k. This together with (43), (45), and the dominated convergence theorem imply that $(2\pi)^{n/2}\hat{L}_k$ converges to the characteristic (indicator) function of Q^n in $L^1(R^n)$ and hence, the desired result follows by taking Fourier transforms. ∎

References

[1] N. Aronszajn, T. M. Creese, and L. J. Lipkin, *Polyharmonic Functions*, Clarendon Press, Oxford, 1983.

[2] H. Cartan, *Elementary Theory of Analytic Functions of One and Several Complex Variables*, Hermann, Paris, 1963.

[3] C. K. Chui, *Vertex Splines*, Lecture notes, Howard University, 1987.

[4] C. de Boor, K. Hollig, and S. Riemenschnieder, Bivariate cardinal interpolation on a three-direction mesh, *Ill. J. Math.*, 29 (1985), 533-566.

[5] W. F. Donoghue, *Distributions and Fourier Transforms*, Academic, New York, 1969.

[6] J. Duchon, Splines minimizing rotation-invariant seminorms in Sobolev spaces, in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller eds., Springer-Verlag, 1977, 85-100.

[7] N. Dyn and D. Levin, Bell shaped basis functions for surface fitting, in *Approximation Theory and Applications*, Z. Ziegler, ed., Academic Press, New York, 1981, 113-129.

[8] R. Franke, A critical comparison of some methods for interpolation of scattered data, *Math. Comp.* 38 (1982), 181-200.

[9] M. Golomb and H. F. Weinberger, Optimal approximation and error bounds, in *On Numerical Approximation*, R. E. Langer, ed., Madison, 1959, 117-190.

[10] R. L. Harder and R. N. Demarais, Interpolation using surface splines, *J.Aircraft*, 9 (1972), 189-191.

[11] L. Hormander, *Linear Partial Differential Operators*, Springer-Verlag, New York, 1969.

[12] L. Hormander, On division of distributions by polynomials, *Ark. Math.* 3 (1958), 555-568.

[13] E. H. Potter, Multivariate Polyharmonic Spline Interpolation, Ph.D. thesis, Iowa State University, 1981.

[14] I. J. Schoenberg, Contributions to the problem of approximation of equidistant data by analytic functions, *Quart. Appl. Math.*, 4 (1946), 45-99.

[15] I. J. Schoenberg, *Cardinal Spline Interpolation*, CBMS Vol. 12, SIAM, Philadelphia, 1973.

[16] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton, Princeton, 1971.

Solutions of Underdetermined Systems of Linear Equations

W. R. Madych*

Abstract

We (i) outline a general framework for generating solutions to underdetermined systems of equations, (ii) review properties of several specific methods, including minimal norm and maximum entropy, (iii) introduce specific alternate methods for generating non-negative solutions, (iv) compare, via systematic numerical examples, the solutions generated by these methods with those generated by the maximum entropy and minimum norm methods, and (v) consider the nature of the positivity constraint by studying a transparent example.

1 Introduction.

This study is an attempt to obtain some understanding of the nature of certain solutions of underdetermined systems of linear equations with a view to their role in the analysis of various practical inverse problems, specifically those associated with image restoration and computed tomography. For example, one of the questions we are interested in is the following: under what conditions, if any, is the so-called maximum entropy solution better than solutions obtained by other methods. To this end we introduce alternate methods for generating non-negative solutions and discuss the results of several systematic numerical experiments.

More specifically we (i) outline a general framework for generating solutions to underdetermined systems of equations, (ii) review properties of several specific methods, including minimal norm and maximum entropy, (iii) introduce specific alternate methods for generating non-negative solutions, (iv) compare, via systematic numerical examples, the solutions generated by these methods with those generated by the maximum entropy and minimum norm methods, and (v) consider the nature of the positivity constraint by studying a transparent example.

*1980 Math. Subject Classification (1985 Revision). 62G05, 65D99. Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

In this paper the basic issue is the non-uniqueness of the solution. We mention but do not specifically address the important issues of noise, ill-conditionedness, and efficiency of the numerical algorithms used to implement the various methods under consideration.

1.1 The Basic Setup.

Consider the system of linear equations

$$(1) \qquad \sum_{j=1}^{n} a_{ij}x_j = b_i\,, \quad i = 1, \ldots, m.$$

Here, we take the field of scalars to be real. Using standard matrix notation (1) may be expressed as a collection of scalar products

$$(2) \qquad \langle a_i, x\rangle = b_i\,, \quad i = 1, \ldots, m,$$

where $a_i = (a_{i1}, \ldots, a_{in})^T$, $x = (x_1, \ldots, x_n)^T$ and $\langle a_i, x\rangle = a_i^T x$; or even more compactly as

$$(3) \qquad Ax = b$$

where $A = (a_{ij})$ is an m by n matrix and $b = (b_1, \ldots, b_m)^T$.

In what follows the matrix A and the data b are assumed to be known. We are interested in the set of solutions to (3), namely the set

$$S_{A,b} = \{x \ : \ Ax = b\}.$$

To avoid complications which are not germain to the questions under consideration here, we always assume that $m < n$ and that A is of full rank. Thus $S_{A,b}$ is not empty; indeed, it is an $n - m$ dimensional affine manifold in the real Euclidean space R^n.

1.2 Motivation.

In the applications we have in mind, namely image restoration, computed tomography, and related inverse problems, A represents the mathematical model or a discrete analogue of the data aquisition scheme and b is the measured data.

In studying such models many considerations must often be taken into account. For example, in certain models of seismic borehole tomography A is rectangular but not of full rank, see [5]; in this case system (3) is both over and underdetermined. Another complication arises in problems of practical interest by virtue of the fact that there is always some degree of uncertainty in taking physical measurements. Thus in many instances (3) is often replaced with

$$(4) \qquad Ax + \epsilon = b$$

where ϵ is a random vector which represents noise and has certain statistical properties, see [1], [8], [9], [11]. In view of the fact that there are standard techniques to handle the above complications we feel that taking into account such considerations here will unnecessarily complicate the discussion and cloud the main issue which is the lack of uniqueness.

Returning to the problem modeled by (3), it is clear that the desired quantity is a solution which, unfortunately, is not uniquely determined by the measured data. In fact, the set of feasible solutions $S_{A,b}$ is very large indeed.

Ideally, if one could constrain the feasible set of solutions appropriately, (3) should contain enough information to uniquely determine the x which gave rise to the data. Of course, the best that one can usually expect is to obtain a reasonable estimator.

We outline some general methods for constraining the feasible set of solutions in section two. In the third section we consider some familiar examples, specifically the minimum norm and maximum entropy methods, and introduce some new methods. Finally, in section four we consider the consequences of certain constraints, particularly the constraint of componentwise positivity; here we also indicate the results of several numerical experiments.

2 Restricting the Feasible Set.

There are many procedures for restricting the solutions of (3). In this paper we consider two general and often related methodologies. These are described below together with a familiar example. Details of how they are related are contained in subsection 2.3.

2.1 Parametrization.

One method is to assume that the solution of (3) is of a certain form. Namely

$$x = F(\xi) \tag{5}$$

where $\xi = (\xi_1, \ldots, \xi_k)$ is contained in a subset of R^k, call it $\mathcal{P}$, and F is a mapping of $\mathcal{P}$ into a subset $\mathcal{M}$ of R^n. Typically the set $\mathcal{P}$ is an open subset of R^k and the mapping F is one to one and continuously differentiable; in this case $\mathcal{M}$ may be viewed as a k dimensional submanifold of R^n. We will always assume that this is the case and, for readily transparent reasons, see (6) below, take $k = m$. Essentially F is a parametrization of the manifold $\mathcal{M}$.

The problem now reduces to finding values of the acceptable parameter ξ such that

$$AF(\xi) = b. \tag{6}$$

If ξ is any solution of (6) then clearly $x = F(\xi)$ is in the intersection of $\mathcal{M}$ and $S_{A,b}$. The main difficulty with this approach in the general case is assuring that the form (5) is such that (6) has a unique solution for every b which may arise in a given application. Furthermore, except for certain examples, it is difficult to determine $\mathcal{M}$, let alone the intersection of $S_{A,b}$ with $\mathcal{M}$.

One classical example where (6) has a unique solution for every b is the case when $\mathcal{P}$ is R^m,

$$x = F(\xi) = \xi_1 a_1 + \ldots + \xi_m a_m = A^T \xi, \tag{7}$$

and $\mathcal{M}$ is the linear subspace generated by $a_1, \ldots, a_m$. In this case the unique solution of (6) is given by

$$\xi = (AA^T)^{-1} b \tag{8}$$

and the corresponding solution x of (3) is given by

$$x = A^T (AA^T)^{-1} b. \tag{9}$$

(We remind the reader that A is assumed to have rank m which implies that AA^T is invertible.) We will return to this important example later.

2.2 Optimization.

The other general approach is to find the minimum (or maximum) of a scalar valued function $f(x)$ defined on a subset $\mathcal{K}$ of R^n subject to the constraints imposed by (2). In other words, find x which satisfies

$$f(x) = \min\{f(y) \ : \ y \in \mathcal{K} \cap S_{A,b}\}. \tag{10}$$

Of course f should be chosen so that the set $\mathcal{K} \cap S_{A,b}$ is not empty and f has a unique minimum on this set. Fortunately by choosing f with certain readily verifiable properties, for instance, convexity, it is not difficult to guarantee that the desired conditions are satisfied.

For example if f is a positive definite quadratic form on $\mathcal{K} = R^n$ then it is a classical fact that (10) has a unique solution, see [10]. In the special case

$$f(x) = \langle x, x \rangle$$

the solution is known as the *minimum norm solution* of (3) and is given by (9).

2.3 A connection.

As mentioned earlier, often the methods of parametrization and optimization are related. To see this, assume that $f(x)$ is continuously differentiable and formally apply the method of Lagrange multipliers to solve problem (10). In

particular, set $\xi = (\xi_1, \dots, \xi_m)^T$ where the ξ_i are the 'Lagrange multipliers' and write

$$h(\xi, x) = f(x) - \sum_{i=1}^{m} \xi_i(\langle a_i, x\rangle - b_i). \tag{11}$$

Taking the gradient of h and setting the result to 0 gives

$$\langle a_i, x\rangle = b_i\ , \quad i = 1, \dots, m,$$

and

$$\frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{n} \xi_i a_{ij} = 0\ ,\ j = 1, \dots, n.$$

The first set of equations is just (2). The last set of equations can be written more compactly and transparently as

$$\nabla f(x) = A^T \xi \tag{12}$$

where ∇f denotes the gradient of f. If ∇f is invertible the optimal solution may be expressed as

$$x = G(A^T \xi) \tag{13}$$

where G is the inverse of ∇f. Finally, equations (2) and (13) imply that under appropriate conditions the solution to problem (10) is given by x of the form (13) where ξ is a solution of

$$AG(A^T \xi) = b. \tag{14}$$

The relationship mentioned above should now be clear. Roughly speaking, if $F(\xi) = G(A^T \xi)$, where G is the inverse of ∇f then the two methods should give rise to the same solution. This observation is useful when trying to determine whether (5) has a unique solution and other questions related to F.

In the case where f is of the form

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \tag{15}$$

equation (13) is particularly simple . Namely, it can be expressed as

$$x_j = g_j((A^T \xi)_j)\ , \quad j = 1, \dots, n, \tag{16}$$

where $(A^T \xi)_j$ denotes the j-th component of $A^T \xi$ and g_j is the (univariate) derivative of f_j. Essentially all the examples below are of this form.

3 Examples.

3.1 Minimum norm and generalizations.

As indicated earlier the minimum norm solution of (3) is the solution to problem (10) in the case

$$f(x) = \langle x, x \rangle \tag{17}$$

Perhaps a more accurate description would be the minimum quadratic norm solution.

Standard variants of (17) are more general quadratics which include f's of the form

$$f(x) = q(C(x-y))$$

where $q(x) = \langle x, x \rangle$, C is an $n \times n$ matrix, and y is a constant vector. The parameters in C and y are usually chosen to influence the behavior of the solution. Properties of such solutions are well known and well documented, for example see [2], [3], [10], and the references cited there.

Perhaps the most important feature of this method is the fact that the relationship between the data and the resulting estimator is linear. Besides being a convenient theoretical tool, this property allows for efficient computational algorithms in many applications.

3.2 Maximum entropy.

The maximum entropy solution of (1) is an estimator whose components are of the form

$$x_j = p_j \exp\left(\sum_{i=1}^{n} \xi_i a_{ij}\right) \ , \ j = 1, \ldots, n, \tag{18}$$

where $p = (p_1, \ldots, p_n)^T$ is a constant and the ξ_i's are parameters chosen so that $x = (x_1, \ldots, x_n)^T$ satisfies (1). Using the notation and terminology of subsection 2.1 this form can be expressed more compactly as

$$x = F(\xi) = P \exp(A^T \xi) \ , \tag{19}$$

where P is the constant diagonal matrix with diagonal p and the exponential is interpreted componentwise, namely, if $y = (y_1, \ldots, y_n)^T$ then $\exp(y) = (\exp(y_1), \ldots, \exp(y_n))^T$.

Observe that in this case the manifold $\mathcal{M}$ is contained in the positive cone

$$R_n^+ = \{x \ : \ x_j > 0 \ , \ j = 1, \ldots, n\}.$$

Thus it should be clear that any estimator of this form will have non-negative components.

The parameter p is chosen to influence the behavior of the solution. In the discussion below, unless indicated otherwise, we always assume the components of p to be one.

Observe that the manifold $\mathcal{M}$ can be described so that in low dimensional examples it is relatively easy to visualize. For example, in the case $m = 2$, $n = 3$, if the last column of A is a linear combination of the first and second with coefficients α_1 and α_2 respectively, then $\mathcal{M}$ may be described by

$$\mathcal{M} = \{x \ : \ x_1 > 0 \ , \ x_2 > 0 \ , \ x_3 = x^{\alpha_1} x^{\alpha_2} \ \}.$$

In the general case, if A is such that the first m columns are linearly independent, and the components of x can always be permuted so that the resulting system of equations has this property, then $\mathcal{M}$ is the intersection of the manifolds $\mathcal{M}_i$, $i = 1, \ldots, n-m$, where

$$\mathcal{M}_i = \{x \ : \ x_1 > 0 \ , \ldots , \ x_m > 0 \ , \ x_{m+i} = x^{\alpha_{i1}} \cdots x^{\alpha_{im}} \ \}.$$

and the α_{ij}'s are appropriate constants.

It is not difficult to see that the resulting estimator may also be viewed as the solution of the minimization problem (10) with

$$f(x) = \sum_{j=1}^{n} x_j \log \left(\frac{x_i}{e p_j} \right) \ , \tag{20}$$

where log is the natural logarithm with base e and the constants p_j are those in (18). As suggested in subsection 2.3, an immediate consequence of this formulation is the fact that if the solution hyperplane $S_{A,b}$ intersects the positive cone R^n_+ then $S_{A,b}$ intersects the manifold parametrized by (19) at exactly one point. In other words, if $R^n_+ \cap S_{A,b}$ is not empty then (3) has a unique solution of the form (19).

Note that the the particular normalization of f in (20) results in $x = p$ as the optimal solution in the case of no contraints.

It should be mentioned that the maximum entropy solution is often viewed as that estimator which maximizes the negative of an expression similar to (20). Thus the term maximum. The rationale behind the term entropy is discussed in [4] and [6].

In view of the fact that much work has centered around this method suprisingly little is known concerning the theoretical properties of the resulting estimators beyond the immediate consequences of the definitions.

One interesting fact concerning such solutions to a very special class of linear systems has been given in [4] and [7]. These systems can be described as follows: Suppose $n = lk$ and write the variable $x = (x_1, \ldots, x_n)$ in a rectangular array as shown.

$$\begin{matrix} x_1 & \cdots & x_k \\ \vdots & & \vdots \\ x_{lk-k+1} & \cdots & x \end{matrix} \tag{21}$$

The system of equations then is simply the collection of row sums and column sums of (21), namely,

$$\begin{aligned} \textstyle\sum_{j=1}^{k} x_{(i-1)k+j} &= r_i \quad , \quad i = 1, \ldots, l, \\ \textstyle\sum_{i=1}^{l} x_{(i-1)k+j} &= c_j \quad , \quad j = 1, \ldots, k, \end{aligned} \tag{22}$$

where each row and column sum is positive and

$$r_1 + \cdots + r_l = 1$$

$$c_1 + \cdots + c_k = 1 \ .$$

For this system the maximum entropy solution is given by

$$x_{(i-1)k+j} = r_i c_j \ . \tag{23}$$

In other words, the value of each component is the product of the row and column sums which contain it.

The setup involving (21), (22), and (23) has the following probabilistic interpretation. If the x_j's represent the probabilities of certain basic events ω_j, $j = 1, \ldots, n$, then the sums in (22) represent the probabilities of the unions,

$$\rho_i = \cup_{j=1}^{k} \omega_{(i-1)k+j}, \ i = 1, \ldots, l \text{ and } \gamma_j = \cup_{i=1}^{l} \omega_{(i-1)k+j}, \ j = 1, \ldots, k.$$

Formula (23) expresses the fact that ρ_i and γ_j are mutually independent events in the probabilistic sense. This interpretation is valid only for systems of the type described by (22); presently there are no analogous results for more general systems.

3.3 Methods for generating bounded solutions.

In this subsection we introduce two alternate methods for generating solutions which are bounded componentwise based on the generalities in the second section.

Observe that it suffices to restrict our attention to those methods which generate estimators which are in the positive cone R_+^n since the change of variables $x \to x - y$ will easily transform such a method to one which generates estimators which are bounded componentwise from below by y. A similar remark holds concerning boundedness from above.

If $\xi = (\xi_1, \ldots, \xi_m)^T$, $-\infty < \xi_i < \infty$, $i = 1, \ldots, m$, consider the parametrization given by

$$x_j = \frac{1}{2}\left\{(A^T\xi)_j + \sqrt{(A^T\xi)_j^2 + 4}\right\} \ , \ j = 1, \ldots, n, \tag{24}$$

where $(A^T\xi)_j$ denotes the j-th component of $A^T\xi$. This parametrization can be expressed more compactly by the formula

$$x = \Phi(A^T\xi)\ , \tag{25}$$

where if $y = (y_1, \ldots, y_n)^T$ then the j-th component of Φ is given by

$$\Phi_j(y) = \frac{1}{2}\left\{y_j + \sqrt{y_j^2 + 4}\right\}\ ,\ j = 1, \ldots, n.$$

Observe that if $\mathcal{M}_\Phi$ denotes the manifold parametrized by (24) then $\mathcal{M}_\Phi$ is contained in the positive cone R_+^n. Thus any solution of (3) which enjoys the representation (24) must have non-negative components.

Proposition 1 *If the intersection of $S_{A,b}$ and R_+^n is not empty then*

$$A\Phi(A^T\xi) = b \tag{26}$$

has a unique solution $\hat{\xi}$. In other words, $\mathcal{M}_\Phi \cap S_{A,b}$ contains exactly one element $\hat{x} = \Phi(A^T\hat{\xi})$.

Proof Consider the scalar function f defined on R_+^n by the formula

$$f(x) = \sum_{j=1}^{n}\left(\frac{x_j^2}{2} - \log x_j\right). \tag{27}$$

It should be clear from the definition of f that there is a unique optimal $\hat{x}$ which minimizes

$$\{f(x)\ :\ x \in S_{A,b} \cap R_+^n\ \}$$

in R_+^n. Using the chain of reasoning outlined in subsection 2.3 it is apparent that such an $\hat{x}$ satisfies

$$\nabla f(\hat{x}) = A^T\hat{\xi}$$

for a unique $\hat{\xi}$. Since Φ is the inverse of ∇f we may write

$$\hat{x} = \Phi(A^T\hat{\xi})$$

and the desired result follows. ∎

For another example consider the relation

$$s = t - \frac{1}{t^3} \tag{28}$$

for positive numbers t. Since the right hand side of (28) is an increasing function of t it is evident that (28) defines a function $\psi(s)$ mapping $0 < s < \infty$ onto $-\infty < t < \infty$. Note that $\psi(s)$ is a root of the polynomial

$t^4 - st^3 - 1$. Now, using roughly the same notation as in (25), we define another parametrization by the formula

$$x = \Psi(A^T \xi)\,, \tag{29}$$

where the j-th component of Ψ is given by

$$\Psi_j(y) = \psi(y_j)\ ,\ j = 1, \ldots, n,$$

and ψ is the function defined earlier in this paragraph.

It is not difficult to see that everything that was said concerning Φ also holds for Ψ. Indeed, we may state the following.

Proposition 2 *Proposition 1 remains true if Φ is replaced by Ψ.*

Proof Recall the proof of Proposition 1. If we simply replace (27) by

$$f(x) = \sum_{j=1}^{n} \frac{1}{2}\left(x_j^2 + \frac{1}{x_j^2}\right) \tag{30}$$

the rest of the proof of this proposition is the same as that of Proposition 1. ■

A direct consequence of the above propositions are two methods for generating estimators for (3). For future reference we will refer to them as method one and method two.

4 Numerical experiments and comments.

Recall that the estimators generated by the method of minimum norm are linear functions of the data. Furthermore there are a host of efficient algorithms for computing them, see [2] and [10]. On the other hand, the estimators generated by the other methods mentioned above depend non-linearly on the data and, as is readily evident, are considerably more difficult to compute. Thus understanding the nature of these estimators and their relative merits is of some practical significance.

4.1 Description of numerical experiments.

To obtain a sense of the nature of the estimators generated by the methods outlined in section 3 we performed numerical experiments on relatively small linear systems. The systems considered were of the form

$$\frac{1}{k}\sum_{j=0}^{k-1} x_{i+j} = b_i\ ,\ i = 1, \ldots, m, \tag{31}$$

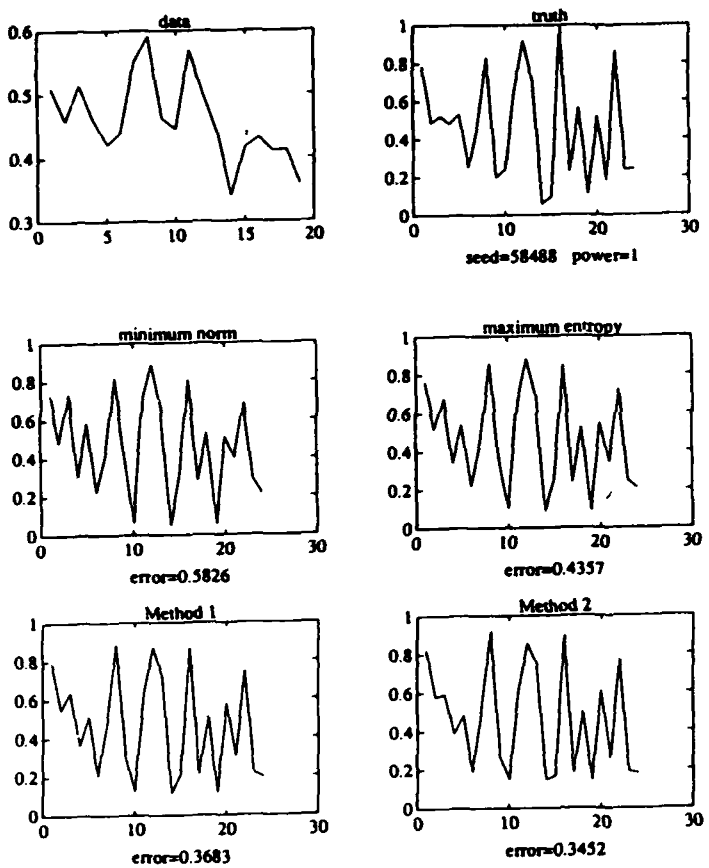
data
truth
seed=58488 power=1
minimum norm
error=0.5826
maximum entropy
error=0.4357
Method 1
error=0.3683
Method 2
error=0.3452

Fig. 1

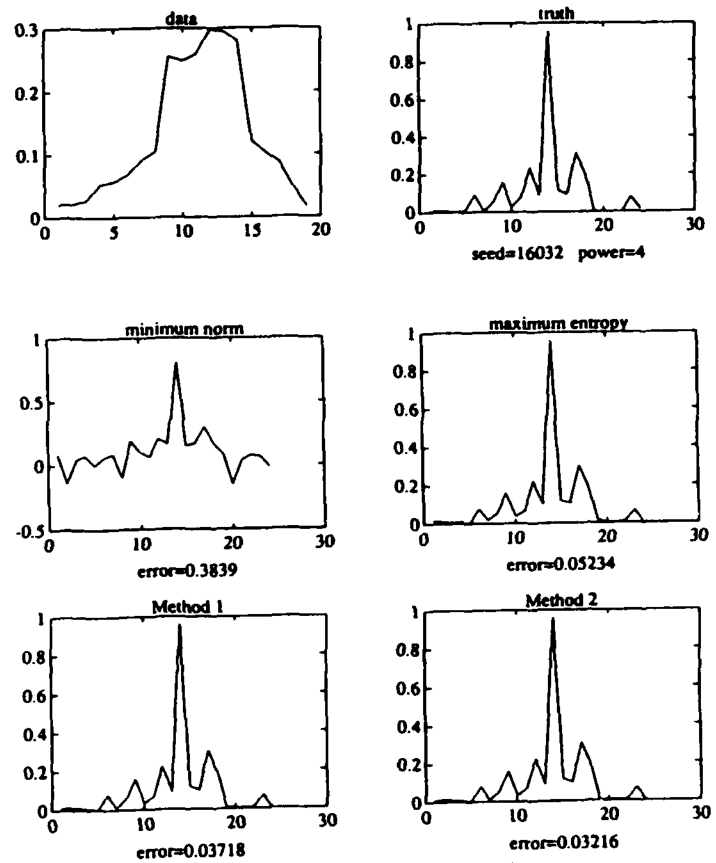
data
truth
seed=16032 power=4
minimum norm
error=0.3839
maximum entropy
error=0.05234
Method 1
error=0.03718
Method 2
error=0.03216

Fig. 2

where $k < n$, $m = n - k + 1$, and n is the number of variables x_j. System (31) is a typical example of a one dimensional blurring model. The sizes considered for our experiments ranged from $m = 15$, $n = 20$ to $m = 80$, $n = 100$ with the ratio m/n varying between 0.5 and 0.9. The experiments were performed as follows:

A non-negative pseudo-random vector $x = (x_1, \ldots, x_n)^T$ was generated via a canned subroutine and the data $b = Ax$ was computed. Then each of the methods outlined in section 3 was used to generate an estimator. We refer to these method as minimum norm, maximum entropy, method one, and method two or, more briefy, MN, ME, M1, and M2 respectively. Formula (9) was used to compute the MN estimator. The other estimators were computed by using Newton's method to solve (6) for $\hat{\xi}$ and then using (5) to evaluate the corresponding estimator $\hat{x}$. In the case of maximum entropy the values $p_j = 1$, $j = 1, \ldots, n$, were used. The resulting estimators were plotted together with the true phantom x. In each case the error

$$\|x - \hat{x}\| = \left\{ \sum_{j=1}^{n} |x_j - \hat{x}_j| \right\}^{1/2}$$

was computed.

The results of these experiments can be loosely described as follows:

When the lower bound, zero in this case, was not a tight one for the phantom then all the methods generated estimates which were roughly equivalent. Namely, they differed from one another but the differences were judged not significant. The computed error varied but was roughly the same order of magnitude for all the methods. For example, see Figure 1; here $m = 19$, $n = 24$ and the components of the phantom x are uniformly distributed between 0 and 1.

On the other hand, when the lower bound for the phantom was reasonably tight then the methods which enforced the lower bound generated considerably better estimators. Indeed, the computed errors for estimators generated by ME, M1, and M2 were significantly smaller than the error for the estimator generated by MN. The estimators generated by ME, M1, and M2 differed but the difference was judged not significant; the same was true of the corresponding computed error. For example, see Figure 2; here again $m = 19$, $n = 24$ but the components of the phantom x are distributed according to the 1/4-th power of the uniform distribution between 0 and 1.

Similar experiments were performed on systems of the form

$$\sum_{j=1}^{n} x_j \cos \frac{2\pi ij}{n}, \quad i = 0, \ldots, k,$$

$$\sum_{j=1}^{n} x_j \sin \frac{2\pi ij}{n}, \quad i = 1, \ldots, k,$$

where $m = 2k + 1$ and m significantly less than n. Here the results were roughly the same as those reported above, although in the cases when the lower bound was tight on the phantom the differences seemed less dramatic.

4.2 Comments.

To obtain some perspective on the observations recorded in the previous subsection consider the system given by

$$x_1 + x_2 = \epsilon \text{ and } x_2 + x_3 = 1 \tag{32}$$

where ϵ is a small positive number. In this case $S_{A,b}$ is simply the line in R^3 parametrized by

$$x = (\epsilon - t, t, 1 - t)\ , \ -\infty < t < \infty.$$

Suppose that, in addition to (32) we know that the phantom which gave rise to the data had non-negative components. Namely,

$$x_1 \geq 0,\ , x_2 \geq 0,\ x_3 \geq 0. \tag{33}$$

It is then clear that with this additional information the feasible set of solutions is that part of $S_{A,b}$ contained in a ball of radius $\sqrt{3}\epsilon/2$ centered at $(\epsilon/2, \epsilon/2, 1 - (\epsilon/2))$. Thus any non-negative solution of (32) will be within $\sqrt{3}\epsilon$ of the desired phantom. Indeed, if $\epsilon = 0$, then (32) together with (33) imply a unique solution.

The principle illustrated by the above simple example is no doubt valid for much larger and more complicated systems of equations. The extent to which this principle is valid depends on the matrix A and the phantom x and should be possible to characterize quantitatively in terms of these parameters.

References

[1] S. Geman and D. E. McClure, "Bayesian image analysis: An application to single photon emission tomography," Proc. American Statistical Association, Statistical Computing Section, (1985), 12-18.

[2] G. Golub and C. Van Loan, Matrix Computations, John Hopkins, Baltimore, 1983.

[3] M. Golomb and H. F. Weinberger, "Optimal Approximation and Error Bounds," in On Numerical Approximation, R. E. Langer, ed. Madison, 1959, 117-190.

[4] S. F. Gull and J. Skilling, "Maximum entropy method in image processing," IEEE Proceedings, 131, (1984), 646-659.

[5] S. Ivanson, "Seismic Borehole Tomography - Theory and Computational Methods," Proc. IEEE, 74, (1986), 328-338.

[6] E. T. Jaynes, Papers on probability statistics and statistical physics, R. Rosencrantz, ed., Reidel, Dordrecht, 1983.

[7] A. K. Livesey and J. Skilling, "Maximum Entropy Theory," Acta Cryst., A41, (1985), 113-122.

[8] W. R. Madych and S. A. Nelson, "Polynomial based algorithms for computed tomography," SIAM J. Applied Math., 43, (1983), 157-185.

[9] L. A. Shepp, and Y. Vardi, "Maximum liklihood reconstruction in positron emission tomography," IEEE Trans. on Medical Imaging, 1, (1982) 113-122.

[10] G. Strang, Introduction to Applied Mathematics, Wellesley-Cambridge, Wellesley, 1986.

[11] D. M. Titterington, "Regularisation procedures in signal processing and statistics," in Mathematics in Signal Processing, T. S. Durrani *et al.* eds., Clarendon, Oxford, 1987, 213-223.

Department of Mathematics, U-9
University of Connecticut
Storrs, CT 06268

Multivariate Interpolation and Conditionally Positive Definite Functions II

W. R. MADYCH* AND S. A. NELSON†

Abstract. We continue an earlier study of certain spaces that provide a variational framework for multivariate interpolation. Using the Fourier transform to analyze these spaces, we obtain error estimates of arbitrarily high order for a class of interpolation methods that includes multiquadrics.

1. Introduction. This paper continues a study, [11], of certain subspaces C_h of $C(\mathbf{R}^n)$, the continuous complex valued functions on n-space $\mathbf{R}^n$. The spaces C_h provide a variational framework for the following interpolation problem: given numerical values at a scattered set of points in $\mathbf{R}^n$, make a good choice of a function f in $C(\mathbf{R}^n)$ that takes on those values.

For the reader's convenience we review some basic features of the development in [11]. The starting point is the selection of an integer $m \geq 0$ and a continuous function h on $\mathbf{R}^n$ that is conditionally positive definite of order m. For example: $m = 1$, $h(x) = -\sqrt{1+|x|^2}$. Using h, a space C_h with a semi-inner product $(\ ,\ )_h$ is constructed. C_h is a subspace of $C(\mathbf{R}^n)$, and the null space of $(\ ,\ )_h$ is P_{m-1}, the polynomials on $\mathbf{R}^n$ of degree $m-1$ or less. A key property of C_h is this: if $x_1, \ldots, x_N$ are distinct points in $\mathbf{R}^n$ and $v_1, \ldots, v_N$ are complex numbers, then among all functions f in C_h that satisfy the interpolation conditions $f(x_i) = v_i$, the quadratic $\|f\|_h^2 = (f, f)_h$ is minimized by a function of the form $f = s + p$ where p is in P_{m-1} and

$$s(x) = \sum_{i=1}^{N} c_i h(x - x_i) \tag{1.1}$$

with $\sum_{i=1}^{N} c_i x_i^\alpha = 0$ for all $|\alpha| < m$. For the example mentioned, (1.1) is a multiquadric interpolant.

Because the spaces C_h are translation invariant, the Fourier transform is a natural tool for analyzing them; it plays a central role here. To clarify basic ideas and make an orderly division of our results, we avoided Fourier techniques in [11]. We did, however, rely on them in our earlier investigation [10], which was in fact prompted by the Fourier methods in Duchon [5]. Use of Fourier transforms allows us to give improved descriptions of the spaces C_h (see Section 3) and allows us to single out certain cases where error estimates of order $\ell \geq m$ are possible (see Section 4). These estimates apply to the multiquadric case as well as to related examples given in Section 5; for each example given there, the integer ℓ can be arbitrarily large.

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Both authors were partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145.
†Department of Mathematics, Iowa State University, Ames, IA 50011.

2. Preliminaries. In this section we recall some notation and results involving Fourier transforms and conditionally positive definite functions.

Let $\mathcal{D}(\mathbf{R}^n)$ denote the space of complex valued functions on $\mathbf{R}^n$ that are compactly supported and infinitely differentiable. The Fourier transform of a function φ in $\mathcal{D}$ is

$$\widehat{\varphi}(\xi) = \int e^{-i(x,\xi)}\varphi(x)dx. \tag{2.1}$$

In order to make use of theorems from Gelfand & Vilenkin [7] we adopt their definition of m-th order conditional positive definiteness. (Equivalence with the definition used in [11] can be seen from Proposition 2.4 and Theorem 6.1 below.) Thus for a continuous function h we assume

$$\int h(x)\varphi * \tilde{\varphi}(x)dx \geq 0 \tag{2.2}$$

holds whenever $\varphi = p(D)\psi$ with ψ in $\mathcal{D}$ and $p(D)$ a linear homogeneous constant coefficient differential operator of order m. Here $\tilde{\varphi}(x) = \overline{\varphi(-x)}$ and $*$ denotes the convolution product

$$\varphi_1 * \varphi_2(t) = \int \varphi_1(x)\varphi_2(t-x)dx.$$

Note that (2.2) can be rewritten as

$$\iint h(x-y)\varphi(x)\overline{\varphi(y)}dxdy \geq 0. \tag{2.3}$$

The following result can be found in Chapter II, Section 4.4 of [7]; we incorporate a remark at the end of that section concerning the case where h is continuous.

Theorem 2.1. *Let h be continuous and conditionally positive definite of order m. Then it is possible to choose a positive Borel measure μ on $\mathbf{R}^n \sim \{0\}$, constants a_γ, $|\gamma| \leq 2m$ and a function χ in $\mathcal{D}$ such that: $1 - \widehat{\chi}(\xi)$ has a zero of order $2m+1$ at $\xi = 0$; both of the integrals $\int_{0<|\xi|<1} |\xi|^{2m} d\mu(\xi)$, $\int_{|\xi|\geq 1} d\mu(\xi)$ are finite; for all $\psi \in \mathcal{D}$,*

$$\int h(x)\psi(x)dx = \int \Big[\widehat{\psi}(\xi) - \widehat{\chi}(\xi) \sum_{|\gamma|<2m} D^\gamma \widehat{\psi}(0)\frac{\xi^\gamma}{\gamma!}\Big] d\mu(\xi) + \sum_{|\gamma|\leq 2m} D^\gamma \widehat{\psi}(0)\frac{a_\gamma}{\gamma!}. \tag{2.4}$$

This uniquely determines the measure μ and the constants a_γ for $|\gamma| = 2m$. In addition, for every choice of complex numbers c_α, $|\alpha| = m$,

$$\sum_{|\alpha|=m}\sum_{|\beta|=m} a_{\alpha+\beta}c_\alpha \bar{c}_\beta \geq 0. \tag{2.5}$$

The choice of χ affects the value of the coefficients a_γ for $|\gamma| < 2m$. Note that the value of the right side of (2.4) does not change if, for suitable φ, $\widehat{\chi}$ is replaced by $\widehat{\chi} + \varphi$ and the a_γ, for $|\gamma| < 2m$, are replaced by $a_\gamma + \int \varphi(\xi)\xi^\gamma d\mu(\xi)$.

As can be seen from

$$(2.6) \qquad (-i)^{|\gamma|} \int x^\gamma \varphi(x)dx = D^\gamma \widehat{\varphi}(0),$$

changing a coefficient a_γ on the right hand side of (2.4) corresponds to changing $h(x)$ on the left side by adding a constant multiple of x^γ.

For $m = 0$, (2.4) reduces to $\int h\psi = \int \widehat{\psi} d\lambda$ where λ is the Borel measure on $\mathbf{R}^n$ given by

$$\lambda(E) = \mu(E \sim \{0\}) + a_0 \delta(E).$$

Here δ is the measure corresponding to a unit mass at the origin; $\delta(E) = 1$ if $0 \in E$ and $\delta(E) = 0$ otherwise. Recall that Borel measures that are finite on compact sets are called Radon measures. We make the usual identification of a Radon measure on an open set $\Omega \subset \mathbf{R}^n$ with the corresponding distribution in $\mathcal{D}'(\Omega)$ and write $\langle \lambda, \psi \rangle = \int \psi d\lambda$. Also, if $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ we identify it with the distribution in $\mathcal{D}'$ given by $\langle f, \psi \rangle = \int \psi(x) f(x) dx$. Thus, for $m = 0$, (2.4) says $\langle h, \varphi \rangle = \langle \lambda, \widehat{\varphi} \rangle$.

For an illustration of the theorem when $m \neq 0$, take $n = 2$, $m = 1$, $h(x) = -\sqrt{1 + |x|^2}$. Then $d\mu(\xi) = w(\xi) d\xi$ with

$$w(\xi) = \frac{(1 + |\xi|)\, e^{-|\xi|}}{(2\pi)^2 |\xi|^3}$$

and $a_\gamma = 0$ for $|\gamma| = 2$. If χ is even, then the coefficients a_γ for $|\gamma| = 1$ are also 0. The remaining coefficient is $a_0 = -\left(1 + \int [1 - \widehat{\chi}(\xi)]\, w(\xi) d\xi\right)$. Details for this and related examples are given in Section 5.

We use $T^k\varphi$ to denote the k-th order Taylor polynomial for φ about 0:

$$(2.7) \qquad T^k \varphi(\xi) = \sum_{|\alpha| \leq k} D^\alpha \varphi(0) \frac{\xi^\alpha}{\alpha!}.$$

The integral on the right side of (2.4) can then be written as $\int \widehat{\psi} - \widehat{\chi} T^{2m-1} \widehat{\psi} d\mu$.

The Schwartz space of rapidly decreasing C^∞ functions and its dual, the space of tempered distributions, are denoted by the usual letters $\mathcal{S}$ and $\mathcal{S}'$.

Proposition 2.2. *Let k be a positive integer and let σ be a Radon measure on $\mathbf{R}^n \sim \{0\}$ such that $\int |\xi|^k (1 + |\xi|^k)^{-1} d|\sigma|(\xi) < \infty$. Let s be a continuous function such that $|\xi|^k s(\xi)$ is bounded on $\mathbf{R}^n$ and $1 - s(\xi) = O\left(|\xi|^k\right)$ at $\xi = 0$. Let*

$$(2.8) \qquad u(x) = \int \left[e^{-i\langle x, \xi \rangle} - s(\xi) \sum_{r=0}^{k-1} \frac{(-i\langle x, \xi \rangle)^r}{r!} \right] d\sigma(\xi).$$

Then $u \in C(\mathbf{R}^n)$, $u(x) = o\left(|x|^k\right)$ as $|x| \to \infty$ and for all φ in $\mathcal{S}$

$$(2.9) \qquad \int u(x) \varphi(x) dx = \int \left(\widehat{\varphi} - s T^{k-1} \widehat{\varphi}\right) d\sigma.$$

PROOF: Let $E(t) = e^{-it} - \sum_{r=0}^{k-1}(-it)^r/r!$ and note that $u = u_0$ where

$$u_a(x) = \int\limits_{|\xi|>a} (1 - s(\xi))\, e^{-i\langle x,\xi\rangle} + s(\xi) E\left(\langle x,\xi\rangle\right) d\sigma(\xi).$$

From $|E(t)| \le |t|^k$ we have $|s(\xi)E(\langle x,\xi\rangle)| \le |x|^k|\xi|^k|s(\xi)|$. Our assumptions on σ and s ensure that $1 - s(\xi)$ and $|\xi|^k|s(\xi)|$ belong to $L^1(\sigma)$. Continuity of u can be established using dominated convergence.

To prove $u(x) = o\left(|x|^k\right)$ note that $|u_0(x) - u_a(x)| \le \left(c_1(a) + c_2(a)|x|^k\right)$ where $c_1(a)$ and $c_2(a)$ are the results of integrating $|1 - s(\xi)|$ and $|\xi|^k|s(\xi)|$ over $0 < |\xi| \le a$ with respect to $|\sigma|$. Given $\varepsilon > 0$, choose $a > 0$ so that $c_1(a) < \varepsilon$ and $c_2(a) < \varepsilon$. From $|E(t)| \le 2|t|^{k-1}$ and $a > 0$ we have $u_a(x) = O\left(|x|^{k-1}\right)$ as $|x| \to \infty$. Thus we may choose $R \ge 1$ such that $|u_a(x)| \le \varepsilon|x|^k$ for all $|x| > R$. Then for $|x| > R$,

$$|u(x)| \le |u_a(x)| + |u_0(x) - u_a(x)| \le \varepsilon|x|^k + \varepsilon + \varepsilon|x|^k.$$

It follows that $u(x) = o\left(|x|^k\right)$.

To establish (2.9), apply Fubini's theorem and use

$$\int \frac{(-i\langle x,\xi\rangle)^r}{r!}\varphi(x)dx = \sum_{|\alpha|=r} D^\alpha\widehat{\varphi}(0)\frac{\xi^\alpha}{\alpha!}.$$

This can be verified by using $(y_1 + \cdots + y_n)^r/r! = \sum\limits_{|\alpha|=r} y^\alpha/\alpha!$ and (2.6). ∎

If u is defined by (2.8) with $\sigma = \mu$, $k = 2m$ and $s = \widehat{\chi}$, then from (2.4), (2.9) and (2.6) we have $\langle h - u, \psi\rangle = \langle q, \psi\rangle$ for all ψ in $\mathcal{D}$. Here $q(x) = \sum\limits_{|\gamma|\le 2m} a_\gamma(-ix)^\gamma/\gamma!$.

Corollary 2.3. *Suppose h is continuous and positive definite of order m. If $m > 0$ then there are unique constants a_γ, $|\gamma| = 2m$ such that $h(x) - \sum_{|\gamma|=2m} a_\gamma(-ix)^\gamma/\gamma! = o\left(|x|^{2m}\right)$, as $|x| \to \infty$. These constants are the same as those appearing in (2.4).*

For ease in dealing with (2.5), we develop some related notation. Let V_m be the space of vectors $v = (v_\alpha)_{|\alpha|=m}$ and let A be the operator on V_m defined by $Av = w$ where $w_\alpha = \sum\limits_{|\beta|=m} A_{\alpha,\beta}v_\beta$ and $A_{\alpha,\beta} = a_{\alpha+\beta}/(\alpha!\beta!)$. Because of (2.5) A must be real symmetric. Thus $Av = 0$ iff $v^T\overline{Av} = 0$. Equivalently, the null space, N_A, of A is the null space of the semi-inner product $(v,w)_A = v^T\overline{Aw}$. Let $H_A = V_m/N_A$ be the Hilbert space obtained by identifying v and w whenever $\|v - w\|_A = 0$. The elements of H_A are the cosets $v + N_A$ and as w varies over such a coset, Aw remains fixed.

By applying Theorem 2.1 we can recover (2.2) for a more convenient set of functions φ. Let

$$\mathcal{D}_m = \left\{\varphi \in \mathcal{D} : \int x^\alpha\varphi(x)dx = 0 \quad \text{for all } |\alpha| < m\right\}. \tag{2.10}$$

Clearly, $\mathcal{D}_m = \{\varphi \in \mathcal{D} : \widehat{\varphi}(\xi) = O(|\xi|^m) \text{ at } \xi = 0\}$. If $\psi = \varphi * \tilde{\varphi}$ then $\widehat{\psi} = |\widehat{\varphi}|^2$ so

$$D^\gamma \widehat{\psi} = \sum_{\alpha+\beta=\gamma} \frac{\gamma!}{\alpha!\beta!} D^\alpha \widehat{\varphi} D^\beta \overline{\widehat{\varphi}}.$$

Hence, for $\psi = \varphi * \tilde{\varphi}$ with $\varphi \in \mathcal{D}_m$,

$$\sum_{|\gamma| \le 2m} D^\gamma \widehat{\psi}(0) \frac{a_\gamma}{\gamma!} = \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha+\beta} \frac{D^\alpha \widehat{\varphi}(0)}{\alpha!} \frac{D^\beta \overline{\widehat{\varphi}(0)}}{\beta!} = \|\widehat{\varphi}^{(m)}(0)\|_A^2 \tag{2.11}$$

where $\widehat{\varphi}^{(m)}(0)$ is the vector v in V_m given by $v_\alpha = D^\alpha \widehat{\varphi}(0)$. From (2.4) we see that if $\varphi \in \mathcal{D}_m$ then

$$\int h(x) \varphi * \tilde{\varphi}(x) dx = \int |\widehat{\varphi}|^2 d\mu + \|\widehat{\varphi}^{(m)}(0)\|_A^2. \tag{2.12}$$

and (2.2) holds. Since $\mathcal{D}_m$ includes the functions φ for which (2.2) was assumed, we conclude that requiring (2.2) for all $\varphi \in \mathcal{D}_m$ is an equivalent definition of h being conditionally positive definite of order m.

Since $\mathcal{D}_{m+1} \subset \mathcal{D}_m$, the latter definition makes it clear that h will be conditionally positive definite of order $m+1$ if it is conditionally positive definite of order m. If m is replaced by $m+1$ in Theorem 2.1, with h held fixed, the measure μ will remain the same, the coefficients a_γ, $|\gamma| = 2(m+1)$ will be 0 and the lower order coefficients will change to reflect changes in $\widehat{\chi}$ and additional terms in the Taylor polynomial.

In order to apply results from [11] we verify that h is in the space $Q_m(\mathbf{R}^n)$ defined there.

Proposition 2.4. *Let h be continuous and assume (2.2) holds for all $\varphi \in \mathcal{D}_m$. If $x_1, \ldots, x_N$ are distinct points in $\mathbf{R}^n$ and $c_1, \ldots, c_N$ are constants that satisfy $\sum_{i=1}^N c_i x_i^\alpha = 0$ for all $|\alpha| < m$, then*

$$\sum_{i,j=1}^N c_i \bar{c}_j h(x_i - x_j) \ge 0. \tag{2.13}$$

Proof: Choose g in $\mathcal{D}$ with $\int g(x)dx = 1$ and $g(x) = 0$ for all $|x| \ge 1$. For $\varepsilon > 0$, let $g_\varepsilon = \varepsilon^{-n} g(x/\varepsilon)$ and take $\varphi_\varepsilon(x) = \sum_{k=1}^N c_k g_\varepsilon(x - x_k)$. Then $\widehat{\varphi_\varepsilon}(\xi) = \tau(\xi)\widehat{g}(\varepsilon\xi)$ with $\tau(\xi) = \sum_{k=1}^N c_k e^{-i\langle x_k, \xi\rangle}$. From $D^\alpha \tau(\xi) = \sum_{k=1}^N c_k (-ix_k)^\alpha e^{-i\langle x_k,\xi\rangle}$ we find $\tau(\xi) = O(|\xi|^m)$ at $\xi = 0$. Thus $\varphi_\varepsilon \in \mathcal{D}_m$ and

$$0 \le \int h(x) \varphi_\varepsilon * \tilde{\varphi}_\varepsilon(x) dx = \iint h(t-y) \varphi_\varepsilon(t) \overline{\varphi_\varepsilon(y)} dt dy.$$

Letting $\varepsilon \to 0$ we obtain (2.13). ∎

The following observations will be used in the next section. Let $\widehat{\mathcal{D}_m} = \{\widehat{\varphi} : \varphi \in \mathcal{D}_m\}$.

Proposition 2.5. *Let $m \geq 0$ and let μ be a positive Borel measure on $\mathbf{R}^n \sim \{0\}$ that satisfies $\int(|\xi|^m/(1+|\xi|^m))^2 d\mu(\xi) < \infty$. If $2k \geq m$ then $\widehat{\mathcal{D}_{2k}}$ is a dense subset of $L^2(\mu)$.*

PROOF: Let $g \in L^2(\mu)$ and $\varepsilon > 0$. Choose $g_1 \in \mathcal{D}(\mathbf{R}^n \sim \{0\})$ so that $\|g - g_1\|_{L^2(\mu)} < \varepsilon$. Then $f(\xi) = |\xi|^{-2k} g_1(\xi)$ is in $\mathcal{D}$. Since $\widehat{\mathcal{D}}$ is dense in $\mathcal{S}$ we can find $\psi \in \mathcal{D}$ so that for all ξ in $\mathbf{R}^n$, $|f(\xi) - \widehat{\psi}(\xi)| \leq \varepsilon/\left(1 + |\xi|^{2k}\right)$. Multiplying by $|\xi|^{2k}$ gives

$$|g_1(\xi) - |\xi|^{2k}\widehat{\psi}| \leq \frac{\varepsilon|\xi|^{2k}}{1+|\xi|^{2k}}.$$

Let $\varphi = (-\Delta)^k \psi$. Then $\varphi \in \mathcal{D}$, $\widehat{\varphi}(\xi) = |\xi|^{2k}\widehat{\psi}(\xi)$ and

$$\int |g_1 - \widehat{\varphi}|^2 d\mu \leq \varepsilon^2 \int \left(\frac{|\xi|^{2k}}{1+|\xi|^{2k}}\right)^2 d\mu(\xi).$$

Thus $\|g - \widehat{\varphi}\|_{L^2(\mu)}$ can be made as small as desired with $\varphi \in \mathcal{D}_{2k}$. ∎

Proposition 2.6. *If $T \in \mathcal{D}'$ satisfies $T(\varphi) = 0$ for all φ in $\mathcal{D}_m$ then T belongs to P_{m-1}.*

PROOF: Define $T_\alpha \in \mathcal{D}'$ by $T_\alpha(\varphi) = \int x^\alpha \varphi(x)dx$ and note that $\bigcap\{T_\alpha^{-1}(0) : |\alpha| < m\} = \mathcal{D}_m$. By assumption, $\mathcal{D}_m$ is contained in $T^{-1}(0)$, the null space of T. It follows (see Theorem 1.3 of [9]) that there are constants c_α such that $T = \sum_{|\alpha|<m} c_\alpha T_\alpha$. ∎

3. Fourier Description of C_h. After analyzing the space $\mathcal{C}_{h,m}$ defined below, we will see that it coincides with the space C_h studied in [11]. Among the results emerging from this analysis is a Fourier transform description of $\mathcal{C}_{h,m}$.

Definition. *Let h be a continuous function on $\mathbf{R}^n$ that is conditionally positive definite of order m. We write $f \in \mathcal{C}_{h,m}(\mathbf{R}^n)$ if $f \in C(\mathbf{R}^n)$ and there is a constant $c(f)$ such that for all φ in $\mathcal{D}_m$*

$$\left|\int f(x)\varphi(x)dx\right| \leq c(f)\left\{\iint h(x-y)\varphi(x)\overline{\varphi(y)}dxdy\right\}^{1/2}. \tag{3.1}$$

If $f \in \mathcal{C}_{h,m}(\mathbf{R}^n)$ we let $c_*(f)$ denote the smallest constant for which (3.1) is true.

It is easily checked that if f_1 and f_2 are in $\mathcal{C}_{h,m}$ then $f_1 + f_2$ and af_1, $a \in \mathbf{C}$, are also in $\mathcal{C}_{h,m}$ with $c_*(f_1+f_2) \leq c_*(f_1) + c_*(f_2)$ and $c_*(af_1) = |a|c_*(f_1)$. If $f \in P_{m-1}$ and $\varphi \in \mathcal{D}_m$ then $\langle f, \varphi\rangle = 0$ so $f \in \mathcal{C}_{h,m}$ and $c_*(f) = 0$. Conversely, if $c_*(f) = 0$ then $f \in P_{m-1}$ by Proposition 2.6. Thus $c_*(f)$ is a semi-norm with null space P_{m-1}; for $m = 0$, take $P_{-1} = \{0\}$.

Using (2.12) we note that (3.1) is equivalent to

$$|\langle f, \varphi\rangle| \leq c(f)\left\{\|\widehat{\varphi}\|^2_{L^2(\mu)} + \|\widehat{\varphi}^{(m)}(0)\|^2_A\right\}^{1/2} \tag{3.2}$$

for all φ in $\mathcal{D}_m$. If $v \in V_m$ and

$$q(x) = \sum_{|\alpha|=m} (Av)_\alpha(-ix)^\alpha \tag{3.3}$$

then $\langle q,\varphi\rangle = \sum_{|\alpha|=m} (Av)_\alpha D^\alpha\widehat{\varphi}(0) = \left(\widehat{\varphi}^{(m)}(0),\bar{v}\right)_A$ so $q \in \mathcal{C}_{h,m}$ with $c_*(q) = \|\bar{v}\|_A$. If $g \in L^2(\mu)$ and u is defined by (2.8) with $\sigma = g\mu$, $k = m$ and an appropriate choice of s (take $s = 0$ for $m = 0$) then, for $\varphi \in \mathcal{D}_m$, (2.9) gives $\langle u,\varphi\rangle = \int \widehat{\varphi} g\, d\mu$. It follows that $u \in \mathcal{C}_{h,m}$ with $c_*(u) = \|g\|_{L^2(\mu)}$.

Clearly $\mathcal{C}_{h,m}$ includes all functions of the form $f = u + q + p$ with u, q as above and $p \in P_{m-1}$. The next result, when combined with Proposition 2.6, shows that all functions in $\mathcal{C}_{h,m}$ can be obtained in this way.

From the behavior of $u(x)$ as $|x| \to \infty$, described by Proposition 2.2, we see that if $m > 0$ and $f = u + q + p$ then $f(x) = o(|x|^m)$ is equivalent to $q = 0$ (or $Av = 0$). In any case

$$\mathcal{C}_{h,m}(\mathbf{R}^n) \subset \left\{f \in C(\mathbf{R}^n) : f(x) = O\left(|x|^m\right) \text{ as } |x| \to \infty\right\}. \tag{3.4}$$

Proposition 3.1. *Let m, h, μ and a_γ be as in Theorem 2.1. If $f \in \mathcal{C}_{h,m}$ then there is a function $g \in L^2(\mu)$ and a vector $v \in V_m$ such that for all φ in $\mathcal{D}_m$*

$$\langle f,\varphi\rangle = \int \widehat{\varphi} g\, d\mu + \sum_{|\alpha|=m} (Av)_\alpha D^\alpha\widehat{\varphi}(0). \tag{3.5}$$

This uniquely determines g and the coset $v + N_A$.

PROOF: Define $J : \mathcal{D}_m \to H = L^2(\mu) \oplus H_A$ by $J\varphi = \widehat{\varphi} \oplus \left(\widehat{\varphi}^{(m)}(0) + N_A\right)$. From (3.2) we see that $|\langle f,\varphi\rangle| \le c_*(f)\|J\varphi\|_H$. From this we deduce that if $J\varphi_1 = J\varphi_2$ then $\langle f,\varphi_1\rangle = \langle f,\varphi_2\rangle$. It follows that there is a bounded linear functional L on the image $J\mathcal{D}_m$, such that $L(J\varphi) = \langle f,\varphi\rangle$ for all φ in $\mathcal{D}_m$. Since H is a Hilbert space, we can choose $\bar{g} \oplus (\bar{v} + N_A)$ so that for all φ in $\mathcal{D}_m$ $\langle f,\varphi\rangle = (J\varphi, \bar{g} \oplus (\bar{v} + N_A))_H$. This gives (3.5).

For uniqueness, we show that $J\mathcal{D}_m$ is dense in H. Let $g_1 \in L^2(\mu)$, $w \in V_m$ and $\eta > 0$ be given. Take $2k > m$ and use Proposition 2.5 to choose $\varphi_1 \in \mathcal{D}_{2k}$ with $\|g_1 - \widehat{\varphi}_1\|_{L^2(\mu)} < \eta$. Note that $J\varphi_1 = \widehat{\varphi}_1 \oplus 0$ since $2k > m$. Put $p(\xi) = \sum_{|\alpha|=m} w_\alpha \xi^\alpha/\alpha!$ and take $\chi \in \mathcal{D}$ so that $1 - \widehat{\chi}(\xi) = O\left(|\xi|^{m+1}\right)$ at $\xi = 0$. Define $\psi_\varepsilon \in \mathcal{D}$ by $\widehat{\psi_\varepsilon}(\xi) = p(\xi)\widehat{\chi}(\varepsilon^{-1}\xi)$. Then $J\psi_\varepsilon = \widehat{\psi_\varepsilon} \oplus (w + N_A)$. Choosing ε close enough to 0, we have $\|\widehat{\psi_\varepsilon}\|_{L^2(\mu)} < \eta$. Then $\|g_1 + (w + N_A) - J(\varphi_1 + \psi_\varepsilon)\|_H < 2\eta$. ∎

If $f \in \mathcal{C}_{h,m}$ let $\Lambda f = g \oplus (v + N_A)$ be the point in $H = L^2(\mu) \oplus H_A$ determined by (3.5). Clearly the resulting map $\Lambda : \mathcal{C}_{h,m} \to H$ is linear. That Λ maps onto H is evident from the remarks leading up to Proposition 3.1. From (3.2) and (3.5) we see that $c_*(f) = \|\Lambda f\|_H$. Note $\|\Lambda f\|_H = \{(f,f)_h\}^{1/2} = \|f\|_h$ where $(f_1,f_2)_h = (\Lambda f_1, \Lambda f_2)_H$ is a semi-inner product for $\mathcal{C}_{h,m}$. There is a corresponding inner product on $\mathcal{C}_{h,m}/P_{m-1}$ which is then a Hilbert space, isomorphic to H under the quotient map associated with Λ.

The following provides a converse to Proposition 3.1 and clarifies how the Fourier transform relates f to g, v in (3.5).

Proposition 3.2. *Let m, h, μ and a_γ be as in Theorem 2.1. Fix $g \in L^2(\mu)$, $v \in V_m$ and $f \in \mathcal{D}'$. The following are equivalent:*

(a) *(3.5) holds for all φ in $\mathcal{D}_m$,*

(b) $f \in S'$ *and for every* $|\alpha| = m$, $\xi^\alpha F = \lambda_\alpha$ *where* F *is the inverse Fourier transform of* f *and* λ_α *is the Radon measure on* $\mathbf{R}^n$ *given by*

$$\lambda_\alpha(E) = \int_{E\sim\{0\}} \xi^\alpha g(\xi)d\mu(\xi) + \alpha!(Av)_\alpha \delta(E). \tag{3.6}$$

When this is the case, $f \in C_{h,m}$, $\Lambda f = g \oplus (v + N_A)$ *and* $(f,f)_h = \int |g|^2 d\mu + v^T \overline{Av}$.

PROOF: Let q be as in (3.3) and let u be defined by (2.8) with $\sigma = g\mu$, $k = m$ and a choice of s that satisfies the hypotheses of Proposition 2.2. If (a) holds then $\langle f, \varphi\rangle = \langle u + q, \varphi\rangle$ for all $\varphi \in \mathcal{D}_m$. By Proposition 2.6 $f - (u+q) = p \in P_{m-1}$. If $\widehat{F} = f$ and $\widehat{\psi}(\xi) = \xi^\alpha \varphi(\xi)$ then

$$\begin{aligned}\langle \xi^\alpha F, \varphi\rangle &= \langle F, \widehat{\psi}\rangle = \langle f, \psi\rangle = \langle u, \psi\rangle + \langle q + p, \psi\rangle \\ &= \int \left(\widehat{\psi} - sT^{m-1}\widehat{\psi}\right) g\, d\mu + \sum_{|\alpha|\le m} b_\alpha D^\alpha \widehat{\psi}(0)\end{aligned}$$

where the constants b_α are determined by $q + p(x) = \sum_{|\alpha|\le m} b_\alpha (ix)^\alpha$. Thus

$$\langle \xi^\alpha F, \varphi\rangle = \int (\xi^\alpha \varphi(\xi) - 0)\, g(\xi)d\mu(\xi) + \alpha!(Av)_\alpha \varphi(0) \tag{3.7}$$

which establishes (b). To see that (b) implies (a), let $f_1 = u + q$ with u and q as above. Then (3.7) holds for F_1 where $\widehat{F}_1 = f_1$. Hence $\xi^\alpha F_1 = \lambda_\alpha$. If (b) holds then $\xi^\alpha F_1 = \xi^\alpha F$ for all $|\alpha| = m$. This implies $F_1 - F = \sum_{|\alpha|<m} b_\alpha D^\alpha \delta$ which says $f_1 - f \in P_{m-1}$. Therefore, (a) and the other assertions about f follow from the corresponding facts about f_1. ∎

For typical choices of h (e.g. those considered in Section 5) the measure μ is absolutely continuous with respect to Lebesgue measure, $d\mu(\xi) = w(\xi)d\xi$, and $a_\gamma = 0$ for all $|\gamma| = 2m$. In such cases the measures λ_α in (3.6) are given by functions F_α in $L^1_{\text{loc}}(\mathbf{R}^n)$; $d\lambda_\alpha(\xi) = F_\alpha(\xi)d\xi$, where $F_\alpha(\xi) = \xi^\alpha g(\xi)w(\xi)$. From $D^\alpha f = ((-i\xi)^\alpha F)^\frown = (-i)^m \widehat{\lambda_\alpha}$, we see that $(D^\alpha f)^\frown = (-i)^m (2\pi)^n \check{F}_\alpha \in L^1_{\text{loc}}(\mathbf{R}^n)$ where $\check{F}_\alpha(\xi) = F_\alpha(-\xi)$. Let

$$r(\xi) = \frac{1}{(2\pi)^{2n}|\xi|^{2m} w(-\xi)} \tag{3.8}$$

with $r(\xi) = \infty$ when $w(-\xi) = 0$. If $d\rho(\xi) = r(\xi)d\xi$, then $(D^\alpha f)^\frown \in L^2(\rho)$ and

$$\|(D^\alpha f)^\frown\|^2_{L^2(\rho)} = \int \frac{\xi^{2\alpha}|g(\xi)|^2}{|\xi|^{2m}} d\mu(\xi).$$

Using (4.2) below with $\ell = m$,

$$\sum_{|\alpha|=m} \frac{m!}{\alpha!} \|(D^\alpha f)^\frown\|^2_{L^2(\rho)} = \int |g|^2 d\mu = (f,f)_h. \tag{3.9}$$

Corollary 3.3. *Let* m, h, μ, *and* a_γ *be as in Theorem 2.1. Assume* $d\mu(\xi) = w(\xi)d\xi$ *and* $a_\gamma = 0$ *for all* $|\gamma| = 2m$. *Let* ρ *be the Borel measure on* $\mathbf{R}^n$ *defined by* $d\rho(\xi) = r(\xi)d\xi$ *with* r *as in (3.8). Then* $f \in C_{h,m}$ *iff* $f \in S'$ *and* $(D^\alpha f)^\frown \in L^2(\rho)$ *for every* $|\alpha| = m$. *In that case* $(f,f)_h$ *is given by (3.9).*

The translation invariant nature of $C_{h,m}$ is evident in the following.

Proposition 3.4. *Let τ be a compactly supported Radon measure on $\mathbf{R}^n$. If f is in $\mathcal{C}_{h,m}$ then so is $\tau * f$. Furthermore, if $\Lambda : \mathcal{C}_{h,m} \to L^2(\mu) \oplus H_A$ is as defined above and $\Lambda f = g \oplus (v + N_A)$ then $\Lambda(\tau * f) = tg \oplus (t(0)v + N_A)$ where $t(\xi) = \int e^{i\langle x,\xi\rangle} d\tau(x)$.*

PROOF: If $\psi(x) = \int \varphi(x+y)d\tau(y)$ then $\langle \tau * f, \varphi\rangle = \langle f, \psi\rangle$ and

$$\widehat{\psi}(\xi) = \iint e^{-i\langle x,\xi\rangle}\varphi(x+y)dx d\tau(y) \qquad (3.10)$$
$$= \iint e^{-i\langle z-y,\xi\rangle}\varphi(z)dz d\tau(y) = \widehat{\varphi}(\xi)t(\xi).$$

If $\Lambda f = g \oplus (v + N_A)$, so that (3.5) holds, then for all $\varphi \in \mathcal{D}_m$

$$\langle \tau * f, \varphi\rangle = \int \widehat{\psi} g\, d\mu + \sum_{|\alpha|=m} D^\alpha \widehat{\psi}(0)(Av)_\alpha$$
$$= \int \widehat{\varphi} tg\, d\mu + \sum_{|\alpha|=\tau} t(0) D^\alpha \widehat{\varphi}(0)(Av)_\alpha.$$

This gives (3.5), with f, g, v replaced by $\tau * f$, tg, $t(0)v$; the assertions made are now apparent. ∎

In the next result, (3.11) is equivalent to $\Lambda(\nu * h) = n \oplus (w + N_A)$ and (3.12) says $\nu(\bar{f}) = (\nu * h, f)_h$. From this it is clear that $\mathcal{C}_{h,m}$ satisfies condition (c) in Theorem 1.1 of [11]. That conditions (a) and (b) are also satisfied can be seen from the discussion above in which the map Λ was introduced. Applying Theorem 1.1 of [11], we conclude that $\mathcal{C}_{h,m} = \mathcal{C}_h$.

Proposition 3.5. *Let m, h, μ and a_γ be as in Theorem 2.1. Let ν be a compactly supported Radon measure on $\mathbf{R}^n$ and assume that $\int x^\alpha d\nu(x) = 0$ for all $|\alpha| < m$. Then $\nu * h \in \mathcal{C}_{h,m}$ and for all φ in $\mathcal{D}_m$*

$$\langle \nu * h, \varphi\rangle = \int \widehat{\varphi} n\, d\mu + \sum_{|\alpha|=m} (Aw)_\alpha D^\alpha \widehat{\varphi}(0) \qquad (3.11)$$

where $n(\xi) = \int e^{i\langle x,\xi\rangle} d\nu(x)$ and $w_\beta = D^\beta n(0) = \int (ix)^\beta d\nu(x)$. Furthermore, if $f \in \mathcal{C}_{h,m}$ and $\Lambda f = g \oplus (v + N_A)$, then

$$\int \overline{f(x)} d\nu(x) = \int n\bar{g}\, d\mu + w^T \overline{Av}. \qquad (3.12)$$

PROOF: If $\psi(z) = \int \varphi(z+y)d\nu(y)$ then, from (2.4),

$$\langle \nu * h, \varphi\rangle = \langle\, , \psi\rangle = \int \widehat{\psi} - \widehat{\chi} T^{2m-1}\widehat{\psi}\, d\mu + \sum_{|\gamma|\le 2m} D^\gamma \widehat{\psi}(0)\frac{a_\gamma}{\gamma!} \qquad (3.13)$$

and, as in (3.10), $\widehat{\psi} = \widehat{\varphi}n$. Clearly, $D^\alpha n(0) = 0$ for all $|\alpha| < m$. If $\varphi \in \mathcal{D}_m$ then $D^\gamma\widehat{\psi}(0) = 0$ for $|\gamma| < 2m$ and for $|\gamma| = 2m$

$$D^\gamma\widehat{\psi}(0) = \sum_{\alpha+\beta=\gamma} \frac{\gamma!}{\alpha!\beta!} D^\alpha\widehat{\varphi}(0)w_\beta.$$

Thus (3.11) follows from (3.13). To establish (3.12) choose a real valued function r in $\mathcal{D}$ with $\widehat{r}(0) = 1$ and for $\varepsilon > 0$ let $\overline{\varphi_\varepsilon(x)} = \int \varepsilon^{-n} r\left(\frac{x-y}{\varepsilon}\right) d\nu(y)$. Then $\varphi_\varepsilon \in \mathcal{D}_m$ and

$$\langle f, \varphi_\varepsilon\rangle = \int \widehat{\varphi_\varepsilon} g\, d\mu + \sum_{|\alpha|=m} (Av)_\alpha D^\alpha \widehat{\varphi_\varepsilon}(0).$$

This yields (3.12) because $\int \overline{f(x)} d\nu(x) = \lim_{\varepsilon\to 0} \overline{\langle f, \varphi_\varepsilon\rangle}$ and $\widehat{\varphi_\varepsilon}(\xi) = \widehat{r}(\varepsilon\xi)\overline{n(\xi)}$. ∎

For s as in (1.1) we have $s = \nu * h$ with $\int \varphi d\nu = \sum_{i=1}^N c_i\varphi(x_i)$. Thus, such functions s belong to $\mathcal{C}_{h,m}$.

The distribution $D^\kappa h$, $|\kappa| \geq m$ can be obtained as a limit of $\nu * h$'s by choosing ν's that correspond to appropriate difference operators. Such ν's satisfy the orthogonality condition $\int x^\alpha d\nu(x) = 0$, $|\alpha| < m$. Hence the following may be regarded as a limiting case of the situation considered above.

Proposition 3.6. *Let m, h, μ and a_γ be as in Theorem 2.1. Fix κ with $|\kappa| \geq m$ and let $p(\xi) = (i\xi)^\kappa$. Then $p \in L^2(\mu)$ iff the distribution $D^\kappa h$ belongs to $\mathcal{C}_{h,m}$. In that case $\Lambda\left((-D)^\kappa h\right) = p \oplus (w + N_A)$ with $w_\alpha = D^\alpha p(0)$, $|\alpha| = m$.*

PROOF: Let $\psi = D^\kappa\varphi$ so $\widehat{\psi} = p\widehat{\varphi}$. If $\varphi \in \mathcal{D}_m$ then, by a calculation like that for (2.11),

$$\sum_{|\gamma|\leq 2m} D^\gamma(p\widehat{\varphi})(0)\frac{a_\gamma}{\gamma!} = \sum_{|\alpha|=m}\sum_{|\beta|=m} a_{\alpha+\beta}\frac{D^\alpha p(0)}{\alpha!}\,\frac{D^\beta\widehat{\varphi}(0)}{\beta!}.$$

Using (2.4), we have

$$\langle(-D)^\kappa h, \varphi\rangle = \langle h, \psi\rangle = \int p\widehat{\varphi} d\mu + \sum_{|\beta|=m} (Aw)_\beta D^\beta\widehat{\varphi}(0) \tag{3.14}$$

for all $\varphi \in \mathcal{D}_m$. This is (3.5) with $f = (-D)^\kappa h$, $g = p$ and $v = w$. If $p \in L^2(\mu)$ we apply Proposition 3.2 to see that $f \in \mathcal{C}_{h,m}$ and $\Lambda f = p \oplus (w + N_A)$. If $p \notin L^2(\mu)$ we apply Proposition 2.5 to obtain a sequence $\varphi_i \in \mathcal{D}_{2k}$ such that $\int |\widehat{\varphi}_i|^2 d\mu = 1$ and $\int p\widehat{\varphi}_i d\mu \to \infty$. We take $2k > m$ so that $D^\beta\widehat{\varphi}_i(0) = 0$ when $|\beta| = m$. Then (3.14) gives

$$\langle(-D)^\kappa h, \varphi_i\rangle = \int p\widehat{\varphi}_i d\mu \to \infty.$$

Since $\|\widehat{\varphi}_i\|^2_{L^2(\mu)} + \|\widehat{\varphi}_i^{(m)}(0)\|^2_A = 1$, we see that $f = (-D)^\kappa h$ cannot satisfy (3.2) and hence cannot be in $\mathcal{C}_{h,m}$. ∎

4. Error Estimates. In this section we derive bounds on the difference between a function g in $\mathcal{C}_{h,m}$ and a function g^X of minimal $\mathcal{C}_{h,m}$ norm that agrees with g on a set $X \subset \mathbf{R}^n$ of 'interpolation points'. These error estimates involve a parameter that measures the spacing of the points in X and are of order ℓ in that parameter; our derivation assumes $\ell \geq m$ and

$$\int |\xi|^{2\ell}\, d\mu(\xi) < \infty. \tag{4.1}$$

For the examples given in Section 5, this assumption is satisfied for arbitrarily large values of ℓ; see (5.2) below. In particular, the estimates apply to multiquadric interpolation since the example there with $a = -1$ gives $h(x) = -2\sqrt{\pi(1+|x|^2)}$.

Before starting on the error estimates, we look at a related implication of (4.1). Let $p_\alpha(\xi) = (i\xi)^\alpha$. From

$$(\xi_1^2 + \cdots + \xi_n^2)^\ell = \sum_{|\alpha|=\ell} \frac{\ell!}{\alpha!}\,\xi^{2\alpha} \tag{4.2}$$

we observe that (4.1) holds iff $p_\alpha \in L^2(\mu)$ for all $|\alpha| = \ell$. If a distribution has all of its ℓ-th order derivatives given by continuous functions then it will belong to $C^\ell(\mathbf{R}^n)$. Thus the following result shows that (4.1) holds iff $\mathcal{C}_{h,m} \subset C^\ell(\mathbf{R}^n)$.

Proposition 4.1. *Let m, h, μ and a_γ be as in Theorem 2.1. Fix α with $|\alpha| \geq m$. Then the following are equivalent:*

(a) *$p_\alpha \in L^2(\mu)$, where $p_\alpha(\xi) = (i\xi)^\alpha$;*

(b) *for every f in $\mathcal{C}_{h,m}$, the distribution $D^\alpha f$ belongs to $C(\mathbf{R}^n)$ and there is a constant c_α such that for all f in $\mathcal{C}_{h,m}$, $\|D^\alpha f\|_\infty \leq c_\alpha \|f\|_h$;*

(c) *there is a point x_0 in $\mathbf{R}^n$ and a constant c_α such that for all f in $\mathcal{C}_{h,m} \cap C^\infty$, $|D^\alpha f(x_0)| \leq c_\alpha \|f\|_h$.*

If these are true, then for all $f \in \mathcal{C}_{h,m}$ and all $y \in \mathbf{R}^n$, $D^\alpha f(y) = (f, \delta_y * (-D)^\alpha h)_h$.

PROOF: Let $f \in \mathcal{C}_{h,m}$ and let F be its inverse Fourier transform so that $\widehat{F} = f$. If $|\alpha| = m$ then, by Proposition 3.2, $\xi^\alpha F = \lambda_\alpha$ with λ_α given by (3.6). If $|\alpha| > m$ then $\alpha = \alpha' + \beta$ with $|\alpha'| = m$. Hence, $\xi^\alpha F = \lambda_\alpha$ with $\lambda_\alpha = \xi^\beta \lambda_{\alpha'}$ where $\lambda_{\alpha'}$ is given by (3.6). If (a) holds then λ_α is finite; for $|\alpha| = m$, $\int d|\lambda_\alpha| = \int |\xi^\alpha g(\xi)| d\mu(\xi) + |(Av)_\alpha|$ and for $|\alpha| > m$, $\int d|\lambda_\alpha| = \int |\xi^\alpha g(\xi)| d\mu(\xi)$. Thus $\widehat{\lambda_\alpha}$ is continuous and bounded by $\int d|\lambda_\alpha|$. Since $(iD)^\alpha f = (\xi^\alpha F)\widehat{} = \widehat{\lambda_\alpha}$ we see that (b) holds with $c_\alpha = \|p_\alpha \oplus (p_\alpha^{(m)}(0) + N_A)\|_H$. Thus (a) implies (b).

That (b) implies (c) is obvious. To see that (c) implies (a), let ψ be an arbitrary function in $\mathcal{D}(\mathbf{R}^n \sim \{0\})$ and define u by (2.8) with $\sigma = \psi\mu$ and $k = m$. Then $u \in \mathcal{C}_{h,m}$, $\Lambda u = \psi \oplus 0$ and $\|u\|_h^2 = \int |\psi|^2 d\mu$. In addition, $u \in C^\infty$ and

$$D^\alpha u(x_0) = \int e^{-i\langle x_0, \xi\rangle}(-i\xi)^\alpha \psi(\xi)\, d\mu(\xi).$$

Thus (c) gives $|\int e^{-i\langle x_0,\xi\rangle}(-i\xi)^\alpha\psi(\xi)d\mu(\xi)| \le c_\alpha\|\psi\|_{L^2(\mu)}$. Since this holds for all ψ in $\mathcal{D}(\mathbf{R}^n \sim \{0\})$, a dense subset of $L^2(\mu)$, (a) must be true.

To verify the last assertion suppose $f \in C_{h,m}$ with $\Lambda f = g \oplus (v + N_A)$. By Proposition 3.6, $\Lambda((-D)^\alpha h) = p_\alpha \oplus (p_\alpha^{(m)}(0) + N_A)$. Using Proposition 3.4 with $\tau = \delta_y$ we have $t(\xi) = e^{i\langle y,\xi\rangle}$ and

$$\Lambda(\delta_y * (-D)^\alpha h) = tp_\alpha \oplus (p_\alpha^{(m)}(0) + N_A). \tag{4.3}$$

Thus $(f, \delta_y * (-D)^\alpha h)_h = \int g\overline{tp_\alpha}d\mu + v^T\overline{Ap_\alpha^{(m)}(0)} = (-i)^m\widehat{\lambda_\alpha}(y)$. Here λ_α is as above so, as already noted, $\widehat{\lambda_\alpha} = (iD)^\alpha\widehat{f}$; this gives the desired equality. ∎

Our error estimates will be based on the following.

Theorem 4.2. *Let m, h, μ and a_γ be as in Theorem 2.1. Assume that μ satisfies (4.1) with $\ell \ge \max\{1, m\}$. For a point x_0 in $\mathbf{R}^n$ suppose that σ is a real valued, compactly supported Radon measure on $\mathbf{R}^n$ such that*

$$p(x_0) = \int p(x)d\sigma(x) \tag{4.4}$$

for all p in $P_{\ell-1}$. Then for all f in $C_{h,m}$

$$\left|f(x_0) - \int f(x)d\sigma(x)\right| \le c\|f\|_h \int |x - x_0|^\ell d|\sigma|(x) \tag{4.5}$$

where $c = \left\{s + \int |\xi|^{2\ell}/(\ell!)^2\, d\mu(\xi)\right\}^{1/2}$ with $s = \sum_{|\alpha|=m}\sum_{|\beta|=m} |A_{\alpha,\beta}|$ for $\ell = m$ and $s = 0$ for $\ell > m$.

Proof: Let $\nu = \delta_{x_0} - \sigma$. By (4.4), $\int p(x)d\nu(x) = 0$ for all $p \in P_{\ell-1}$. Since $\ell \ge m$, Proposition 3.5 applies to ν and from (3.12),

$$\left|\int \overline{f(x)}d\nu(x)\right| \le \|n \oplus (w + N_A)\|_H\|f\|_h. \tag{4.6}$$

Here $w_\beta = \int(ix)^\beta d\nu(x)$, $|\beta| = m$. If $\ell > m$ then $w = 0$; if $\ell = m$ then

$$w_\beta = i^m \int (x - x_0)^\beta d\nu(x) = 0 - i^m \int (x - x_0)^\beta d\sigma(x).$$

Defining $R(\theta)$ by $e^{i\theta} = \sum_{k=0}^{\ell-1} (i\theta)^k/k! + R(\theta)$ we have $|R(\theta)| \le |\theta|^\ell/\ell!$ and

$$e^{-i\langle x_0,\xi\rangle}n(\xi) = \int e^{i\langle x-x_0,\xi\rangle}d\nu(x) = \int R(\langle x - x_0, \xi\rangle)\, d\nu(x) = -\int R(\langle x - x_0, \xi\rangle)\, d\sigma(x).$$

If $b = \int |x - x_0|^\ell d|\sigma|(x)$, then $|n(\xi)| \le b|\xi|^\ell/\ell!$ and, for $\ell = m$, $|w_\beta| \le b$. From this we obtain $\|n \oplus (w + N_A)\|_H \le cb$ and (4.5) follows. ∎

To obtain the error estimates mentioned at the beginning of this section we apply Theorem 4.2 to $f = g - g^X$. Because of the minimum norm property of g^X, $\|f\|_h \leq \|g\|_h$. Since other fixed bounds on $\|f\|_h$ result in acceptable error estimates, the minimum norm requirement on g^X could be relaxed to simply a requirement that $\|g^X\|_h$ not exceed some set bound. If we choose σ so that supp $\sigma \subset X$ then $\int g - g^X d\sigma = 0$ and (4.5) gives

$$|g(x_0) - g^X(x_0)| \leq c\|f\|_h \int |x - x_0|^\ell d|\sigma|(x). \tag{4.7}$$

To make such a choice of σ possible, it may be necessary to restrict x_0. From (4.4) we see that if $p \equiv 0$ on supp σ then $p(x_0) = 0$. Let

$$\begin{aligned} N_{\ell-1}(X) &= \{p \in P_{\ell-1} : p(x) = 0 \quad \text{for all } x \in X\}, \\ \langle X \rangle_{\ell-1} &= \{x \in \mathbf{R}^n : p(x) = 0 \quad \text{for all } p \in N_{\ell-1}(X)\}. \end{aligned}$$

Proposition 4.3. *Let $E_{\ell-1}(x_0, X)$ be the set of all real valued, compactly supported Radon measures on $\mathbf{R}^n$ that satisfy both (4.4) and supp $\sigma \subset X$. Then $E_{\ell-1}(x_0, X)$ is nonempty iff $x_0 \in \langle X \rangle_{\ell-1}$.*

PROOF: Necessity of $x_0 \in \langle X \rangle_{\ell-1}$ is evident from the preceding discussion. To see that this is also sufficient, consider the linear functionals on $P_{\ell-1}$ defined by $L_x(r) = p(x)$. Choose a (finite) subset Y of X such that $\{L_y : y \in Y\}$ is linearly independent and $L_x \in \text{span}\{L_y : y \in Y\}$ for all x in X. Then $N_{\ell-1}(Y) = N_{\ell-1}(X)$ and $\langle Y \rangle_{\ell-1} = \langle X \rangle_{\ell-1}$. Also, $\{L_y : y \in Y\}$ is a basis for $(P_{\ell-1}/N_{\ell-1}(Y))'$; let $\{p_y + N_{\ell-1}(Y) : y \in Y\}$ be the dual basis. If the polynomials p_y are replaced by their real parts the result is still dual to $\{L_y : y \in Y\}$. We may therefore assume that each p_y is real valued. For x_0 in $\langle Y \rangle_{\ell-1}$, L_{x_0} gives a linear functional on $P_{\ell-1}/N_{\ell-1}(Y)$. Thus $L_{x_0} = \sum_{y \in Y} c_y L_y$ with $c_y = L_{x_0}(p_y)$ and it follows that $\sigma = \sum_{y \in Y} c_y \delta_y$ is in $E_{\ell-1}(x_0, X)$. ∎

Of course (4.7) will give a better error estimate if σ is chosen from $E_{\ell-1}(x_0, X)$ so as to minimize $\int |x - x_0|^\ell d|\sigma|(x)$; we made no attempt to do this with our choice of σ in the preceding proof.

We turn now to an analysis of the rate at which the error estimate goes to zero as the coverage by X improves. For this we fix a region Ω and a function $g \in C_{h,m}$ and, for various X, look at bounds on $|g - g^X|_\Omega$ given by (4.7). Here we use the notation $|f|_\Omega = \sup_{x \in \Omega} |f(x)|$.

The number $d = d(\Omega, X)$ defined by

$$d(\Omega, X) = \sup_{y \in \Omega} \inf_{x \in X} |y - x| \tag{4.8}$$

is a standard measurement of how closely X covers Ω. Using (4.7) and some mild assumptions about Ω, we will show that

$$|g - g^X|_\Omega = O(d^\ell). \tag{4.9}$$

In order to use (4.7) we assume (4.1). In that case Proposition 4.1 assures us of a uniform bound, for the ℓ-th order derivatives of $g - g^X$. From this and (4.9) we can deduce that the derivatives $D^\alpha(g - g^X)$ of intermediate order $0 < |\alpha| < \ell$ satisfy $O(d^{\ell-|\alpha|})$ estimates.

To establish (4.9) we proceed along lines used by Duchon [6]. We start by assuming that there are positive constants M, ε_0 such that for every $0 < \varepsilon < \varepsilon_0$,

$$\Omega \subset \bigcup\{B(t, \varepsilon M) : t \in T_\varepsilon\} \tag{4.10}$$

where $T_\varepsilon = \{t \in \mathbf{R}^n : B(t,\varepsilon) \subset \Omega\}$, $B(t,r) = \{x \in \mathbf{R}^n : |x - t| \leq r\}$. Arguments in Section 1 of [6] show that such constants M, ε_0 will exist if Ω satisfies a cone condition.

Next we select a $P_{\ell-1}$-unisolvent set of points $\mathbf{a}(\alpha) \in \mathbf{R}^n$, $|\alpha| < \ell$. A corresponding set of Lagrange polynominals, $p_\gamma^{\mathbf{a}} \in P_{\ell-1}$, $|\gamma| < \ell$, is determined by the requirements: $p_\gamma^{\mathbf{a}}(\mathbf{a}(\alpha)) = 1$, for $\alpha = \gamma$; $p_\gamma^{\mathbf{a}}(\mathbf{a}(\alpha)) = 0$, for $\alpha \neq \gamma$. The matrix $A_{\alpha,\beta} = (\mathbf{a}(\alpha))^\beta$, $|\alpha| < \ell$, $|\beta| <$ is nonsingular. If $p_\gamma(x) = \sum_{|\beta|<\ell} (A^{-1})_{\beta,\gamma} x^\beta$ then $p_\gamma(\mathbf{a}(\alpha)) = (AA^{-1})_{\alpha,\gamma}$ so $p_\gamma = p_\gamma^{\mathbf{a}}$. The function $\alpha \to \mathbf{a}(\alpha)$ can be identified with a point in $\mathbf{B} = \prod_{|\alpha|<\ell} B(\mathbf{a}(\alpha), \delta)$. Clearly, $\mathbf{b} \in \mathbf{B}$ iff $|\mathbf{b}(\alpha) - \mathbf{a}(\alpha)| < \delta$ for all $|\alpha| < \ell$. Now choose $\delta > 0$ so that $B_{\alpha,\beta} = (\mathbf{b}(\alpha))^\beta$ is invertible for all $\mathbf{b} \in \mathbf{B}$. As justified by replacing the points $\mathbf{a}(\alpha)$ with the points $\delta^{-1}\mathbf{a}(\alpha)$, we assume $\delta = 1$.

Choose R so that $B(0, R)$ contains all the unit balls $B(\mathbf{a}(\alpha), 1)$, $|\alpha| < \ell$. The Lagrange polynomials $p_\alpha^{\mathbf{b}}$ depend continuously on $\mathbf{b}$. Let

$$\lambda(r) = \sup\Big\{ \sum_{|\alpha|<\ell} |p_\alpha^{\mathbf{b}}(x)| : |x| \leq r, \mathbf{b} \in \mathbf{B}\Big\}.$$

For $d = d(\Omega, X) < \varepsilon_0/R$, set $\varepsilon = Rd$ and fix a point t in T_ε. The balls $B(t + d\mathbf{a}(\alpha), d)$ are contained in $B(t, Rd) = B(t, \varepsilon) \subset \Omega$. By (4.8), for every $|\alpha| < \ell$, there is at least one point x_α in $X \cap B(t + d\mathbf{a}(\alpha), d)$. If $\mathbf{b}$ is the point in $\mathbf{B}$ defined by $x_\alpha = t + d\mathbf{b}(\alpha)$, and

$$\sigma = \sum_{|\alpha|<\ell} p_\alpha^{\mathbf{b}}\left(\frac{x_0 - t}{d}\right) \delta_{x_\alpha}$$

with x_0 arbitrary, then supp $\sigma \subset X \cap B(t, \varepsilon)$ and (4.4) holds for all $p \in P_{\ell-1}$; to verify (4.4), take q so that $p(x) = q((x - t)/d)$ and use $\sum_{|\alpha|<\ell} p_\alpha^{\mathbf{b}}(y) q(\mathbf{b}(\alpha)) = q(y)$ with $y = (x_0 - t)/d$.

Suppose $x_0 \in B(t, \varepsilon M + d)$. Then $|x_0 - t|/d \leq (RM + 1)$ so $\int d|\sigma| \leq \lambda(RM + 1)$. Also, for $x \in$ supp σ,

$$|x - x_0| \leq |x - t| + |t - x_0| \leq (R + RM + 1)d.$$

Thus $\int |x - x_0|^\ell d|\sigma| \leq C^\circ d^\ell$ with $C^\circ = (R + RM + 1)^\ell \lambda(RM + 1)$. Since x_0 is any point in $B(t, \varepsilon M + d)$, (4.7) gives $|g - g^X|_{B(t,\varepsilon M+d)} \leq c\|f\|_h C^\circ d^\ell$. By (4.10), if $y \in \Omega$ we can choose $t \in T_\varepsilon$ so that $y \in B(t, \varepsilon M)$. Then $B(y, d) \subset B(t, \varepsilon M + d)$ so for every $y \in \Omega$,

$$|g - g^X|_{B(y,d)} \leq cC^\circ \|f\|_h d^\ell. \tag{4.11}$$

This is more than required for (4.9), but will be useful for derivative estimates.

By Proposition 4.1, $f = g - g^X$ is in $C^\ell(\mathbf{R}^n)$. For $y \in \Omega$, $\theta \in \mathbf{R}$ and $u \in \mathbf{R}^n$ with $|u| = 1$, let $\varphi(\theta) = f(y + \theta u)$. Then

$$\varphi^{(k)}(\theta) = k! \sum_{|\alpha|=k} \frac{u^\alpha}{\alpha!} D^\alpha f(y + \theta u). \tag{4.12}$$

By (b) in Proposition (4.1), $|\varphi^{(\ell)}|_{\mathbf{R}} \le C' \|f\|_h$ with $C' = \ell! \sum_{|\alpha|=\ell} c_\alpha/\alpha!$. From (4.11) we also have a bound on $|\varphi|_I$ where I is the interval $[-d, d]$. For $0 < k < \ell$, the results of Gorny [8] summarized in [12] then give

$$|\varphi^{(k)}(0)| \le C_k \|f\|_h d^{\ell-k} \tag{4.13}$$

where $C_k = 16(2e)^k (cC^\circ)^{1-k/\ell} \left[\max(C', \ell! 2^{-\ell} c C^\circ)\right]^{k/\ell}$. Note that C_k can be calculated from n, ℓ, m, h and M; the choice of R depends only on ℓ and n so C° requires only ℓ, n, M while c and C' require only m, h, ℓ, n. Combining (4.12) and (4.13) gives

$$\sup_{|u|=1} \left| \sum_{|\alpha|=k} \frac{u^\alpha}{\alpha!} D^\alpha f(y) \right| \le \frac{C_k}{k!} \|f\|_h \, d^{\ell-k} \tag{4.14}$$

for every $y \in \Omega$. Since

$$|v|_k = \sup_{|u|=1} \left| \sum_{|\alpha|=k} \frac{u^\alpha}{\alpha!} v_\alpha \right|$$

is a norm for V_k, we conclude that $|D^\alpha f|_\Omega = O\left(d^{\ell-|\alpha|}\right)$ for every $|\alpha| \le \ell$. To summarize we state

Theorem 4.4. *Let m, h, μ and a_γ be as in Theorem 2.1. Assume (4.1) holds with $\ell \ge \max\{1, m\}$ and suppose Ω is a subset of $\mathbf{R}^n$ that satisfies (4.10) for some $M, \varepsilon_0 > 0$. Then there are positive constants C, d_0 such that if $f \in \mathcal{C}_{h,m}$ vanishes on a set X and the number $d = d(\Omega, X)$ defined by (4.8) is less than d_0 then for all $|\alpha| \le \ell$,*

$$|D^\alpha f|_\Omega \le C \|f\|_h \, d^{\ell-|\alpha|}. \tag{4.15}$$

5. Examples. In this section we look at some examples of conditionally positive definite functions h. For these examples we determine the measure μ and coefficients $a_\gamma, |\gamma| = 2m$ that appear in (2.4). As can be seen from (5.2) below, these examples all satisfy (4.1) and do so for arbitrarily large choices of ℓ. Thus the error estimates in Section 4 apply, showing that for interpolation based on any of the h's given here, approximation of arbitrarily high order can be achieved.

For $a \in \mathbf{R}$, let w_a be the function on $\mathbf{R}^n$ defined by

$$w_a(\xi) = \frac{2\, K_{(n-a)/2}(|\xi|)}{(2\pi)^{n/2}\, 2^{a/2}\, |\xi|^{(n-a)/2}} \tag{5.1}$$

where K_ν is a modified Bessel function of the third kind. From the behavior of $K_\nu(r)$ at $r = 0$ and $r = \infty$ we note that

(5.2)
$$\int |\xi|^{2\ell} w_a(\xi) d\xi < \infty$$

iff $a + 2\ell > 0$. For $a \in \mathbf{R}$, $a \neq 0, -2, -4, \dots$ let

(5.3)
$$h_a(x) = \frac{\Gamma(a/2)}{(1+|x|^2)^{a/2}},$$

and for $a = -2k, k = 0, 1, 2, \dots$ define h_a by

(5.4)
$$\begin{aligned} h_{-2k}(x) &= \lim_{a \to -2k} \left[h_a(x) - \Gamma(a/2)(1+|x|^2)^k\right] \\ &= \frac{(-1)^{k+1}}{k!}(1+|x|^2)^k \log(1+|x|^2). \end{aligned}$$

The last equality can be verified by using $\Gamma(\frac{a}{2}+k+1) = \left(\frac{a}{2}+k\right)\dots\left(\frac{a}{2}\right)\Gamma\left(\frac{a}{2}\right)$ together with

$$\left.\frac{d}{dt}\right|_{t=k} (1+|x|^2)^t = \lim_{a \to -2k} \frac{(1+|x|^2)^{-a/2} - (1+|x|^2)^k}{(-a/2) - k}.$$

Lemma 5.1. *If $\widehat{\varphi} \in \mathcal{D}(\mathbf{R}^n \sim \{0\})$ then for all a in $\mathbf{R}$*

(5.5)
$$\int h_a(x)\varphi(x)dx = \int \widehat{\varphi}(\xi) w_a(\xi) d\xi.$$

PROOF: A basic fact used in the theory of Bessel potentials is that (5.5) holds for all $\varphi \in \mathcal{S}$ if $a > 0$; see [2], [3] or [4]. For $\widehat{\varphi} \in \mathcal{D}(\mathbf{R}^n \sim \{0\})$ an analytic continuation argument gives (5.5) for $a \neq 0, -2, -4, \dots$. To obtain (5.5) for the remaining values of $a = -2k$, we take limits. If $f(t) = (1+|x|^2)^t$ and $a \neq 0, -2, -4, \dots$, then

$$\left[h_a(x) - \Gamma\left(\frac{a}{2}\right)(1+|x|^2)^k\right] = \left(\frac{a}{2}+k\right)\Gamma\left(\frac{a}{2}\right)\int_0^1 f'\left(k - (\frac{a}{2}+k)s\right) ds.$$

Estimates from thi an be used to justify an application of Lebesgue's dominated convergence theorem that shows

$$\int h_{-2k}(x)\varphi(x)dx = \lim_{a \to -2k} \int \left[h_a(x) - \Gamma\left(\frac{a}{2}\right)(1+|x|^2)^k\right]\varphi(x)dx.$$

Now $\widehat{\varphi} \in \mathcal{D}(\mathbf{R}^n \sim \{0\})$ so $\int (1+|x|^2)^k \varphi(x)dx = 0$. We therefore have $\int h_{-2k}(x)\varphi(x)dx = \lim_{a \to -2k} \int \widehat{\varphi}(\xi) w_a(\xi) d\xi$ which gives (5.5) for $a = -2k$. ∎

Theorem 5.2. *If m is a nonnegative integer and $a+2m>0$ then (2.4) holds with $h=h_a$, $d\mu(\xi)=w_a(\xi)d\xi$, and $a_\gamma=0$ for $|\gamma|=2m$.*

PROOF: If $m=0$ then $a>0$. As already mentioned, (5.5) holds for all φ in $\mathcal{S}$ if $a>0$; thus we have (2.4) with $m=0$ and $a>0$. For the rest of the proof we assume $m\geq 1$. Let

$$u_a(x)=\int\Big[e^{-i\langle x,\xi\rangle}-\widehat{\chi}(\xi)\sum_{k=0}^{2m-1}\frac{(-i\langle x,\xi\rangle)^k}{k!}\Big]w_a(\xi)d\xi.$$

By Proposition 2.2 we have: $u_a\in C(\mathbf{R}^n)$, $u_a(x)=o(|x|^{2m})$, and for all φ in $\mathcal{S}$

$$\int u_a(x)\varphi(x)dx=\langle S_a,\widehat{\varphi}\rangle$$

where $\langle S_a,\psi\rangle=\int\left[\psi-\widehat{\chi}T^{2m-1}\psi\right](\xi)w_a(\xi)d\xi$. Let T_a be the tempered distribution defined by $\int h_a(x)\varphi(x)dx=\langle T_a,\widehat{\varphi}\rangle$. By (5.5), $\langle T_a,\psi\rangle=\langle S_a,\psi\rangle$ for all $\psi\in\mathcal{D}(\mathbf{R}^n\sim\{0\})$. Thus $(T_a-S_a)^\frown=h_a-u_a$ is a polynomial q. Both h_a and u_a are $o(|x|^{2m})$ at $|x|=\infty$, so $\deg q<2m$. The desired instance of (2.4) now follows from $\langle h_a-q,\varphi\rangle=\langle S_a,\widehat{\varphi}\rangle$. ∎

6. Equivalence of definitions. Theorem 6.1 below, when combined with Proposition 2.4 shows the equivalence of the definition of conditional positive definiteness adopted here with that used in [11]. As in [11] we define $P_{m-1}^{\perp}$ to be the space of all finite measures ν on $\mathbf{R}^n$ that have support consisting of a finite set of points and satisfy $\nu(p)=0$ for all $p\in P_{m-1}$. The space obtained by relaxing the support requirement to allow compact sets, rather than only finite sets, will be denoted by $\langle P_{m-1}^{\perp}\rangle$. If $\nu=\sum_{i=1}^N c_i\delta_{x_i}$ then

$$\nu\left(\overline{\nu*\bar{h}}\right)=\sum_{i=1}^N\sum_{j=1}^N c_i\bar{c}_jh(x_i-x_j)$$

and $\nu\in P_{m-1}^{\perp}$ iff $\sum_{i=1}^N c_ix_i^\alpha=0$ for all $|\alpha|<m$. If $d\nu(x)=\varphi(x)dx$ then

$$\nu\left(\overline{\nu*\bar{h}}\right)=\iint\varphi(x)\overline{\varphi(y)}h(x-y)dxdy$$

and ν is in $\langle P_{m-1}^{\perp}\rangle$ if $\varphi\in\mathcal{D}_m$.

Theorem 6.1. *Let h be an arbitrary function in $C(\mathbf{R}^n)$. If $\nu\left(\overline{\nu*\bar{h}}\right)\geq 0$ holds for all $\nu\in P_{m-1}^{\perp}$ then it holds for all $\nu\in\langle P_{m-1}^{\perp}\rangle$.*

PROOF: Fix ν in $\langle P_{m-1}^{\perp}\rangle$ and let K be its support. Recall that the finite Borel measures on K form the dual $C(K)'$ of $C(K)$, the continuous functions on K with the sup norm topology. The norms involved in this duality will be written as follows: for $f\in C(K)$, $|f|_K=\sup_{x\in K}|f(x)|$; for $\sigma\in C(K)'$, $\|\sigma\|=\int d|\sigma|$. Let $h_y(x)=h(y-x)$. K is compact, so for every $\varepsilon>0$ there is a finite set $F_\varepsilon\subset K$ such that if $y\in K$ then $|h_y-h_{y_0}|_K<\varepsilon$ for at least one $y_0\in F_\varepsilon$. If σ is in the weak* neighborhood

$$U(\nu,F_\varepsilon,\varepsilon)=\left\{\sigma\in C(K)':\left|(\sigma-\nu)\left(\bar{h}_{y_0}\right)\right|<\varepsilon\text{ for all }y_0\in F_\varepsilon\right\}$$

and $y \in K$ then, for a suitable choice of $y_0 \in F_\varepsilon$,

$$\left|(\sigma-\nu)\left(\bar{h}_y\right)\right| = \left|(\sigma-\nu)\left(\bar{h}_y - \bar{h}_{y_0}\right) + (\sigma-\nu)\left(\bar{h}_{y_0}\right)\right| \leq (\|\sigma-\nu\|+1)\,\varepsilon.$$

Since $(\sigma-\nu)*\bar{h}(y) = (\sigma-\nu)\left(\bar{h}_y\right)$, we get $|(\sigma-\nu)*\bar{h}|_K \leq (\|\sigma-\nu\|+1)\varepsilon$ for all $\sigma \in U(\nu, F_\varepsilon, \varepsilon)$. For such σ let w be the number defined by

$$w = \sigma\left(\overline{\sigma * \bar{h}}\right) - \nu\left(\overline{\nu * \bar{h}}\right) = \sigma\left(\overline{(\sigma-\nu) * \bar{h}}\right) + (\sigma-\nu)\left(\overline{\nu * \bar{h}}\right)$$

and observe $|w| \leq \|\sigma\| |(\sigma-\nu)*\bar{h}|_K + |(\sigma-\nu)\left(\overline{\nu * \bar{h}}\right)|$.

Let $B = \{\sigma \in C(K)' : \|\sigma\| \leq \|\nu\|\}$ and take $C = B \cap \langle P_{m-1}^{\perp}\rangle, S = B \cap P_{m-1}^{\perp}$. By arguments given below, S is weak* dense in C. This allows us to choose

$$\sigma \in S \cap \{\sigma \in U(\nu, F_\varepsilon, \varepsilon) : |(\sigma-\nu)\left(\overline{\nu * \bar{h}}\right)| < \varepsilon\}.$$

For that choice we have $\sigma\left(\overline{\sigma * \bar{h}}\right) \geq 0$ and

$$|w| \leq \|\sigma\|\,(\|\sigma-\nu\|+1)\,\varepsilon + \varepsilon \leq \|\nu\|\,(2\|\nu\|+1)\,\varepsilon + \varepsilon.$$

Since w is arbitrarily small, we see that $\nu\left(\overline{\nu * \bar{h}}\right)$ must be arbitrarily close to points on the positive real axis and hence must be greater than or equal to zero.

C is convex and weak* compact so, by the Krein-Milman theorem, C is the closed convex hull of its extreme points. Since S is convex it will be weak* dense if it contains all of the extreme points of C. Suppose σ_0 is an extreme point of C that is not in S. Then supp σ_0 cannot be a finite set so we can subdivide it into $J = 2(1 + \dim P_{m-1})$ disjoint subsets $E_1, \ldots, E_J$ with $|\sigma_0|(E_j) \neq 0$. Let $\sigma_j(E) = \sigma_0(E_j \cap E)$ and take $c_{\alpha,j} = \int x^\alpha d\sigma_j(x)$. By a dimension argument, there is a point $a \in \mathbf{R}^J \sim \{0\}$ that satisfies the equations:

$$\sum_{j=1}^{J} a_j \|\sigma_j\| = 0; \quad \sum_{j=1}^{J} a_j c_{\alpha,j} = 0,\ |\alpha| < m.$$

For $t \in \mathbf{R}$, let $\sigma^t = \sum_{j=1}^{J}(1 + t\,a_j)\sigma_j$. Then $\sigma^t \in \langle P_{m-1}^{\perp}\rangle$ and if $(1 + t\,a_j) \geq 0$,

$$\|\sigma^t\| = \sum_{j=1}^{J}(1 + t\,a_j)\|\sigma_j\| = \sum_{j=1}^{J}\|\sigma_j\| = \|\sigma_0\| \leq \|\nu\|.$$

Thus $\sigma^t \in C$ for all t in an interval about 0. This contradicts the assumption that σ_0 was an extreme point of C because $\sigma^t = \sigma_0$ only if $t = 0$, as seen from the fact that $a \neq 0$ and $\|\sigma_j\| \neq 0$ for all $j = 1, \ldots, J$. ∎

REFERENCES

[1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, The Theory of Splines and Their Applications, Academic Press, New York, 1967.

[2] N. Aronszajn and K. T. Smith, *Theory of Bessel potentials. I*, Ann. Inst. Fourier **11** (1961), 385–475.

[3] A. P. Calderon, *Lebesgue spaces of differentiable functions and distributions*, Proc. Symp. in Pure Math. **5** (1961), 33–49.

[4] W. F. Donoghue, Distributions and Fourier Transforms, Academic Press, New York, 1969.

[5] J. Duchon, *Splines minimizing · otation-invariant semi-norms in Sobolev spaces*, pp. 85–100 in Constructive Theory of Functions of Several Variables, W. Schempp and K. Zeller, Ed., Springer-Verlag, Berlin-Heidelberg, 1977.

[6] J. Duchon, *Sur l'erreur d'interpolation des fonctions de plusieurs variables pàr les D^m-splines*, RAIRO Analyse Numerique **12** (1978), 325–334.

[7] I. M. Gelfand and N. Ya. Vilenkin, Generalized Functions, Vol. 4, Academic Press, New York, 1964.

[8] A. Gorny, *Contribution à l'étude des fonctions dérivables d'une variable réelle*, Acta Math. **71** (1939), 317-358.

[9] J. L. Kelly and I. Namioka, Linear Topological Spaces, D. Van Nostrand, Princeton, 1963.

[10] W. R. Madych and S. A. Nelson, *Multivariate interpolation: a variational theory*, manuscript, 1983.

[11] W. R. Madych and S. A. Nelson, *Multivariate interpolation and conditionally positive definite functions*, Approx. Theory and its Applications **4** (1988), no. 4, 77-89.

[12] D. S. Mitrinović, Analytic Inequalities, Springer-Verlag, Berlin-Heidelberg, 1970.

Polyharmonic Cardinal Splines: A Minimization Property

W. R. Madych* S. A. Nelson†

Abstract

Polyharmonic cardinal splines are distributions which are annihilated by iterates of the Laplacian in the complement of a lattice in Euclidean n-space and satisfy certain continuity conditions. Some of the basic properties were recorded in our earlier paper on the subject. Here we show that such splines solve a variational problem analogous to the univariate case considered by I. J. Schoenberg.

1 Introduction

Recall that a k-harmonic cardinal spline is a tempered distribution u on R^n which satisfies

$$(1) \qquad \begin{array}{ll} (i) & u \text{ is in } C^{2k-n-1}(R^n) \text{ and} \\ (ii) & \Delta^k u = 0 \text{ on } R^n \setminus Z^n. \end{array}$$

Here Δ is the usual Laplace operator defined by

$$\Delta u = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2}$$

and, if k is greater than one, Δ^k denotes its k-th iterate, $\Delta^k u = \Delta(\Delta^{k-1} u)$. Of course $\Delta^1 = \Delta$ and Z^n denotes the lattice of points in R^n all of whose coordinates are integers.

Such distributions were considered in [1] where their basic properties were recorded and our motivation for studying them was indicated. One of the key developments there was the existence and uniqueness of the solution to the so-called cardinal interpolation problem for k-harmonic splines. This result may be summarized as follows: Given a sequence $\{a_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, of polynomial growth and an integer k satisfying $2k \geq n+1$ there is a unique k-harmonic spline f such that $f(\mathbf{j}) = a_{\mathbf{j}}$ for all $\mathbf{j}$ in Z^n.

In this paper we continue recording properties of these distributions. Specifically we show that under appropriate conditions the k-harmonic splines are solutions of a variational problem. These properties together with those considered in [1] are remarkably similar to well known properties of the univariate cardinal splines of I. J. Schoenberg, see [2]. For example, much of the material in this paper parallels matter found in [2, Chapter 6]. On the other hand, because of the non-existence of B-splines with compact support in the general

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Both authors were partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

†Department of Mathematics, Iowa State University, Ames, IA 50011.

multivariate case, our development is significantly different from that found there.

The variational problem alluded to above is considered in the context of the space $L_k^2(R^n)$, the class of those tempered distributions whose derivatives of order k are square integrable. The properties of this class and its discrete analogue which are needed for our development are presented in Section 2. In Section 3 it is shown that the class of k-harmonic splines in $L_k^2(R^n)$ is a closed subspace of $L_k^2(R^n)$ whose corresponding orthogonal projection operator is quite natural; this is the key to what may be called the minimization property of these splines. Necessary and sufficient conditions on a sequence $\{v_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, which allow it to be interpolated by the elements of $L_k^2(R^n)$ are given in Section 3; furthermore, it is shown that the unique element of minimal $L_k^2(R^n)$ norm which interpolates such a sequence is the k-harmonic spline interpolant.

The conventions and notation used here are identical to that in [1]. In particular, $SH_k(R^n)$ denotes the space of k-harmonic splines on R^n; k is always assumed to be an integer such that $2k \geq n+1$. The distributions L_k and Φ_k are defined by the formulas for their Fourier transforms,

$$\hat{L}_k(\xi) = (2\pi)^{-n/2} \frac{|\xi|^{-2k}}{\sum_{\mathbf{j}\in Z^n} |\xi - 2\pi\mathbf{j}|^{-2k}} \tag{2}$$

and

$$\hat{\Phi}_k(\xi) = |\xi|^{2k} \hat{L}_k(\xi). \tag{3}$$

Their properties which are relevant to our development are listed in [1]. Here we merely recall that L_k is called the fundamental function of interpolation; it is the unique k-harmonic spline such that $L_k(\mathbf{j}) = \delta_{0\mathbf{j}}$, $\mathbf{j}$ in Z^n, where $\delta_{0\mathbf{j}}$ is the Kronecker delta. In particular, every k-harmonic spline u enjoys the representation

$$u(x) = \sum_{\mathbf{j}\in Z^n} u(\mathbf{j}) L_k(x - \mathbf{j}) \tag{4}$$

where the series converges absolutely and uniformly on compact subsets of R^n. For more details, background, and references see [1].

2 Definition and properties of $L_k^2(R^n)$ and $\ell_k^2(Z^n)$

The linear space $L_k^2(R^n)$ is defined as the class of those tempered distributions u on R^n all of whose k-th order derivatives are square integrable; in other words

$$L_k^2(R^n) = \{u \in \mathcal{S}'(R^n) \,:\, D^\nu u \text{ is in } L^2(R^n) \text{ for all } \nu \text{ with } |\nu| = k\}.$$

For this space a semi-inner product is given by

$$\langle u, v\rangle_k = \sum_{|\nu|=k} c_\nu \int_{R^n} D^\nu u(x) \overline{D^\nu v(x)} dx \tag{5}$$

where the positive constants c_ν are specified by

$$|\xi|^{2k} = \sum_{|\nu|=k} c_\nu \xi^{2\nu}. \tag{6}$$

The semi-norm corresponding to (5) is denoted $\|u\|_{2,k}$, thus $\|u\|_{2,k}^2 = \langle u, u\rangle_k$. The null space of this semi-norm is $\pi_{k-1}(R^n)$, the class of polynomials of degree less than or equal to $k-1$.

Note that there are many semi-norms equivalent to $||u||_{2,k}$ on $L_k^2(R^n)$. The reason for the particular choice used here is the fact that for u in $S(R^n)$, by virtue of Plancherel's formula,

$$||u||_{2,k}^2 = \int_{R^n} |\xi|^{2k}|\hat{u}(\xi)|^2 d\xi.$$

In view of (3) and (4), it is not difficult to conjecture that it is the minimization of this particular semi-norm subject to the appropriate interpolatory conditions which leads to a solution which is a k-harmonic spline.

The objective of this section is to develop properties of $L_k^2(R^n)$ and its discrete analogue which will be needed in our treatment of the variational problem in the following sections.

Since the norm on $L_k^2(R^n)$ only allows us to distinguish between the equivalence classes determined by it, dealing with the individual elements of $L_k^2(R^n)$ is rather slippery business. Nevertheless we regard the space $L_k^2(R^n)$ as a subspace of tempered distribution and not as a collection of equivalence classes. Thus if a distribution u is representable by a continuous function f, namely

$$\langle u, \phi\rangle = \langle f, \phi\rangle = \int_{R^n} f(x)\phi(x)dx$$

for all ϕ in $S(R^n)$, then, following standard convention, we simply identify u with f and say that u is continuous. The following propositions allow us to get a somewhat better grip on the distributions in $L_k^2(R^n)$.

The first proposition follows from a routine argument and may be regarded as folklore.

Proposition 1 *$S(R^n)$ is dense in $L_k^2(R^n)$.*

The notion of unisolvent set is needed in the statement of the next proposition. Recall that a *unisolvent set* Ω for $\pi_{k-1}(R^n)$ is a finite subset of R^n consisting of $k(n)$ elements with the property that if p is in $\pi_{k-1}(R^n)$and $p(x) = 0$ for all x in Ω then $p(x) = 0$ for all x in R^n. Here $\pi_{k-1}(R^n)$ denotes the class of polynomials on R^n of degree no greater than $k-1$ and $k(n)$ denotes its dimension.

Proposition 2 *Assume $2k \geq n+1$. Then the elements of $L_k^2(R^n)$ are continuous functions and there is a linear map P on $L_k^2(R^n)$ with the following properties:*

(i) Pu is in $\pi_{k-1}(R^n)$.

(ii) $P^2u = Pu$.

(iii) If Ω is a unisolvent set for $\pi_{k-1}(R^n)$ then

$$|Pu(x)| \leq C(1+|x|^{k-1})(||u||_{2,k} + ||u||_\Omega) \tag{7}$$

where $||u||_\Omega$ denotes the maximum of u on Ω and C is a constant which depends on Ω but is independent of u.

(iv) If $Qu = u - Pu$ then Qu is continuous and satisfies

$$||Qu||_{2,k} = ||u||_{2,k} \tag{8}$$

and

(9) $$|Qu(x)| \le C(1+|x|^k)\|u\|_{2,k}$$

where C is a constant independent of u.

Proof The operator P is defined by the formula for the Fourier transform of Pu. Namely, if ϕ is any element of $\mathcal{S}(R^n)$ then

(10) $$\langle \widehat{Pu}, \phi\rangle = \langle \hat{u}, p_\phi \chi\rangle$$

where p_ϕ is the Taylor polynomial of ϕ of degree $k-1$ and centered at 0, χ is a fixed function in $C_0^\infty(R^n)$ which is equal to one in a neighborhood of 0 and supported in the unit ball centered at 0. Of course $p_\phi\chi$ denotes the pointwise multiplication of p_ϕ and χ. (Note that the definition of P depends on the choice of χ.) Pu is well defined by formula (10). It is easy to check that P is linear and satisfies statements *(i)* and *(ii)* of the proposition. Next we verify *(iv)*.

Write

(11) $$Qu = u - Pu = u_1 + u_2$$

where $\hat{u}_1 = \chi\widehat{Qu}$, $u_2 = Qu - u_1$, and χ is the function in the definition of P above. More specifically, for any ϕ in $\mathcal{S}(R^n)$

(12) $$\langle \hat{u}_1, \phi\rangle = \langle \hat{u}, (\phi - p_\phi)\chi\rangle$$

and

(13) $$\langle \hat{u}_2, \phi\rangle = \langle \hat{u}, (1-\chi)\phi\rangle.$$

Since $\hat{u}_1$ has compact support, u_1 is analytic and

(14) $$u_1(x) = (2\pi)^{-n/2}\langle \hat{u}_1, e_x\rangle$$

where $e_x(\xi)$ denotes the exponential $e^{i\langle x,\xi\rangle}$. Using (12) write

(15) $$\langle \hat{u}_1, e_x\rangle = \langle \hat{u}, q_x\chi\rangle$$

where

(16) $$q_x(\xi)\chi(\xi) = [e_x(\xi) - p_{e_x}(\xi)]\chi(\xi) = \langle x,\xi\rangle^k \psi_x(\xi)\chi(\xi)$$

and ψ_x is analytic and bounded independent of x. Since

$$\langle x,\xi\rangle^k = \sum_{|\nu|=k} b_\nu x^\nu \xi^\nu$$

formulas (15) and (16) result in

(17) $$\langle \hat{u}_1, e_x\rangle = \sum_{|\nu|=k} b_\nu x^\nu \langle \hat{u}_\nu, \psi_x\chi\rangle$$

where $\hat{u}_\nu = \xi^\nu \hat{u}(\xi)$. Recalling the fact that if $|\nu| = k$ then $\hat{u}_\nu$ is in $L^2(R^n)$ with norm dominated by $\|u\|_{2,k}$ we see that (17) implies that

(18) $$|\langle \hat{u}_1, e_x\rangle| \le C|x|^k\|u\|_{2,k}$$

which together with (14) shows that

(19) $$|u_1(x)| \le C|x|^k\|u\|_{2,k}.$$

From (13) $\hat{u}_2 = (1-\chi)\hat{u}$ and thus $\hat{u}_2$ is in $L^1(R^n)$ since

$$\int_{R^n} (1-\chi(\xi))|\hat{u}(\xi)|d\xi \leq C\left\{\int_{R^n}(1-\chi(\xi))^2|\xi|^{-2k}d\xi\right\}^{1/2}\|u\|_{2,k}.$$

This implies that u_2 is continuous and

$$|u_2(x)| \leq C\|u\|_{2,k}. \tag{20}$$

Now, u_1 and u_2 are both continuous and hence it follows that Qu is also. Identity (11) together with inequalities (19) and (20) imply (9). Since (8) is an immediate consequence of the fact that $u - Qu$ is in $\pi_{k-1}(R^n)$ the proof of statement *(iv)* is complete.

Finally, to see *(iii)* let Ω be the collection of points $\{x_1,\ldots,x_N\}$ where $N = k(n)$ is the dimension of $\pi_{k-1}(R^n)$ and let p_j, $j = 1,\ldots,N$ be the polynomials in $\pi_{k-1}(R^n)$ which are uniquely defined by $p_j(x_m) = \delta_{jm}$ where δ_{jm} is the Kronecker delta. Since $Pu = u - Qu$ is in $\pi_{k-1}(R^n)$ and Ω is unisolvent for $\pi_{k-1}(R^n)$ we see that

$$Pu(x) = \sum_{j=1}^{N}\{u(x_j) - Qu(x_j)\}p_j(x). \tag{21}$$

Now

$$|u(x_j) - Qu(x_j)| \leq |u(x_j)| + C(1+|x_j|^k)\|u\|_{2,k} \tag{22}$$

and

$$|p_j(x)| \leq C_j(1+|x|^{k-1}). \tag{23}$$

Formula (21) together with inequalities (22) and (23) imply the desired result. ∎

The operators P and Q are complementary orthogonal projections on $L_k^2(R^n)$. That is, every u in $L_k^2(R^n)$ can be expressed as

$$u = Pu + Qu. \tag{24}$$

It is clear from the proof that the operator P is not unique.

We now turn our attention to a discrete analogue of $L_k^2(R^n)$ which we refer to as $\ell_k^2(Z^n)$. This class of sequences may be described as follows.

First, recall that $\mathcal{Y}^\alpha$, α real, is the class of those sequences $u = \{u_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, for which the norm

$$N_\alpha(u) = \sup_{\mathbf{j}\in Z^n}\frac{|u_{\mathbf{j}}|}{(1+|\mathbf{j}|)^\alpha} \tag{25}$$

is finite. Let

$$\mathcal{Y}^\infty = \bigcup_\alpha \mathcal{Y}^\alpha \text{ and } \mathcal{Y}^{-\infty} = \bigcap_\alpha \mathcal{Y}^\alpha \tag{26}$$

where the intersection and union are taken over all α, $-\infty < \alpha < \infty$. Also, we say that the sequence $u = \{u_{\mathbf{j}}\}$ is in $\pi_{k-1}(Z^n)$ if it is the restriction of an element p in $\pi_{k-1}(R^n)$ to Z^n; namely, $u_{\mathbf{j}} = p(\mathbf{j})$ for all $\mathbf{j}$ in Z^n where p is in $\pi_{k-1}(Z^n)$.

Given a sequence u in $\mathcal{Y}^{\infty}$, for each i, $i = 1, \ldots, n$, $T_i u$ is the sequence

$$(T_i u)_{\mathbf{j}} = u_{\mathbf{j}+\mathbf{e}_i} - u_{\mathbf{j}}$$

where $\mathbf{e}_i$ is the n-tuple with 1 in the i-th slot and 0 elsewhere. In other words, T_i is a difference operator in the direction $\mathbf{e}_i$. For any multi-index ν T^{ν} is the usual composition (product) of $T_1, \ldots, T_n$, namely $T_1^{\nu_1} \cdots T_n^{\nu_n}$.

The space $\ell_k^2(Z^n)$ consists of those elements $u = \{u_{\mathbf{j}}\}$ in $\mathcal{Y}^{\infty}$ for which

$$\|u\|_{2,k}^2 = \sum_{\mathbf{j} \in Z^n} \sum_{|\nu| = k} |(T^{\nu} u)_{\mathbf{j}}|^2 \tag{27}$$

is finite. The corresponding semi-inner product is given by

$$\langle u, v \rangle_k = \sum_{|\nu| = k} \sum_{\mathbf{j} \in Z^n} (T^{\nu} u)_{\mathbf{j}} \overline{(T^{\nu} v)_{\mathbf{j}}}. \tag{28}$$

The notation used to denote discrete sequences and certain discrete sequence norms is identical to that used for the 'analog' case considered earlier in this section. This should cause no confusion; the meaning should be clear from the context.

Proposition 3 *$\mathcal{Y}^{-\infty}$ is dense in $\ell_k^2(Z^n)$.*

Proposition 4 *There is a linear map P on $\ell_k^2(Z^n)$ with the following properties:*

(i) Pu is in $\pi_{k-1}(Z^n)$.

(ii) $P^2 u = Pu$.

(iii) If Ω is a unisolvent set for $\pi_{k-1}(Z^n)$ then

$$|Pu_{\mathbf{j}}| \le C(1 + |\mathbf{j}|^{k-1})(\|u\|_{2,k} + \|u\|_{\Omega}) \tag{29}$$

where $\|u\|_{\Omega}$ denotes the maximum of u on Ω and C is a constant which depends on Ω but is independent of u.

(iv) If $Qu = u - Pu$

$$\|Qu\|_{2,k} = \|u\|_{2,k} \tag{30}$$

and

$$|Qu_{\mathbf{j}}| \le C(1 + |\mathbf{j}|^{k})\|u\|_{2,k} \tag{31}$$

where C is a constant independent of u.

The proof of these propositions essentially consists of identifying the space $\ell_k^2(Z^n)$ with an appropriate class of tempered distributions and applying the arguments used in proving the analogous facts for $L_k^2(R^n)$ in Propositions 1 and 2 *mutatis mutandis*. Indeed, observe that $\mathcal{Y}^{-\infty}$ equipped with the seminorms defined by (25) is a topological vector space whose dual can be identified with $\mathcal{Y}^{\infty}$. Now, the Fourier transform can be defined on $\mathcal{Y}^{-\infty}$ in the natural way; namely

$$\hat{u}(\xi) = (2\pi)^{-n/2} \sum_{\mathbf{j} \in Z^n} u(\mathbf{j}) e^{-i\langle \mathbf{j}, \xi \rangle}.$$

It maps $\mathcal{Y}^{-\infty}$ into the class of infinitely differentiable periodic functions on R^n. Hence the Fourier transform may be defined on $\mathcal{Y}^{\infty}$ via duality in the usual manner; it maps $\mathcal{Y}^{\infty}$ onto the class of periodic tempered distributions. With these identifications, it should be clear how to modify the arguments used in the proof of Propositions 1 and 2 so that they apply to $\ell^2_k(Z^n)$.

Another very useful identification may be described as follows.

Let $\mathcal{M}(Z^n)$ be the class of those tempered distributions f which enjoy the representation

$$f(x) = \sum_{\mathbf{j} \in Z^n} a_{\mathbf{j}} \delta(x - \mathbf{j}) \tag{32}$$

where $\delta(x)$ denotes the unit Dirac distribution at the origin. Note that $\mathcal{M}(Z^n)$ is a closed subspace of $\mathcal{S}'(R^n)$. Also recall that the Fourier transform is an isomorphism of $\mathcal{M}(Z^n)$ onto the class of periodic distributions.

Observe that $\mathcal{Y}^{\infty}$ and $\mathcal{M}(Z^n)$ are algebraically isomorphic via the mapping

$$u \longrightarrow \sum u_{\mathbf{j}} \delta(x - \mathbf{j}). \tag{33}$$

Thus, whenever convenient, we may view elements of $\mathcal{Y}^{\infty}$ as being in $\mathcal{M}(Z^n)$ and vice versa.

Finite differences, such as those used in the definition of $\ell^2_k(Z^n)$, may be considered elements of $\mathcal{M}(Z^n)$ as follows.

Let $\mathcal{F}$ be the subset of $\mathcal{M}(Z^n)$ consisting of those elements whose representation (32) contains at most a finite number of non-zero coefficients $a_{\mathbf{j}}$. The Fourier transform is an isomorphism of $\mathcal{F}$ onto the space $\mathcal{T}$ consisting of trigonometric polynomials. Note that $\mathcal{F}$ can be identified with a class of finite difference operators via convolution in the natural way; namely, if

$$T(x) = \sum a_{\mathbf{j}} \delta(x - \mathbf{j}) \tag{34}$$

is an element of $\mathcal{F}$ and u is any tempered distribution then

$$T * u(x) = \sum a_{\mathbf{j}} u(x - \mathbf{j})$$

is a finite difference of f. Thus we will often refer to elements of $\mathcal{F}$ as finite differences.

In view of the identification (33), we may view $\mathcal{F}$ as the class of finite difference operators on $\mathcal{Y}^{\infty}$ or $\mathcal{S}'(R^n)$. In particular, the operators T^{ν} used in the definition of $\ell^2_k(Z^n)$ may be regarded as elements of $\mathcal{F}$ whose Fourier transforms are the trigonometric polynomials

$$\widehat{T^{\nu}} = (e^{i\xi_1} - 1)^{\nu_1} \cdots (e^{i\xi_n} - 1)^{\nu_n}. \tag{35}$$

More generally, if T is a finite difference operator of form (34), we write Tu to denote $T * u$ if u is in $\mathcal{S}'(R^n)$ and, if u is in $\mathcal{Y}^{\infty}$, to denote the sequence representing the natural action of T on u, namely,

$$(Tu)_{\mathbf{j}} = \sum_{\mathbf{i}} a_{\mathbf{i}} u_{\mathbf{j}-\mathbf{i}} \, .$$

3 Polyharmonic splines and $L_k^2(R^n)$

Recall that a continuous function f on R^n interpolates a sequence $u = \{u_j\}$ if $f(j) = u_j$ for all j in Z^n. Proposition 2 together with existence and uniqueness for the cardinal interpolation problem for k-harmonic splines, see [1], clearly imply the following.

Proposition 5 *If $2k \geq n+1$ and u is in $L_k^2(R^n)$ then u is continuous and of polynomial growth. The sequence of values $\{u(j)\}$, j in Z^n, is in $\mathcal{Y}^k$ and there is a unique k-harmonic spline which interpolates this sequence.*

If u is in $L_k^2(R^n)$ let $S_k u$ be the unique k-harmonic spline which interpolates the data sequence $\{u(j)\}$, j in Z^n; that is $S_k u(j) = u(j)$ for all j. Recall that we may write

$$S_k u(x) = \sum_{j \in Z^n} u(j) L_k(x - j) \tag{36}$$

where L_k is the fundamental spline defined in the Introduction. Clearly the mapping $u \to S_k u$ is linear and $S_k u = u$ whenever u is in $SH_k(R^n)$. In what follows we will show that this mapping is an orthogonal projection of $L_k^2(R^n)$ onto $SH_k(R^n) \cap L_k^2(R^n)$.

Throughout the rest of this section P and Q will denote a fixed pair of complementary projections whose existence is guaranteed by Proposition 2. We begin with some technical lemmas.

Lemma 6 *Suppose $2k \geq n+1$ and $\{u_m\}$, $m = 1, \ldots,$ is a sequence in $L_k^2(R^n)$. If $\{u_m\}$ converges to u in $L_k^2(R^n)$ then*

(i) $\{Qu_m\}$ converges to Qu uniformly on compact subsets of R^n.

(ii) $\{S_k Qu_m\}$ converges to $S_k Qu$ uniformly on compact subsets of R^n.

Proof Choose any positive numbers r and ϵ and observe that *(i)* will follow if we can show that

$$|Qu_m(x) - Qu(x)| < \epsilon \tag{37}$$

whenever m is sufficiently large and for all x such that $|x| < r$. From (9) we have

$$|Qu_m(x) - Qu(x)| \leq C(1 + |x|^k)\|u_m - u\|_{2,k}. \tag{38}$$

Now from (38) it is easy to see that choosing m so that

$$\|u_m - u\|_{2,k} < \epsilon\{C(1 + r^k)\}^{-1}$$

implies (37).

To see *(ii)* observe that (36), (38), and the linearity of S_k imply

$$|S_k Qu_m(x) - S_k Qu(x)| \leq C \left\{ \sum_{j \in Z^n} (1 + |j|^k) L_k(x - j) \right\} \|u_m - u\|_{2,k}.$$

Now, using the exponential decay of L_k, see [1], it is clear that *(ii)* follows from essentially the same reasoning as *(i)*. ■

Lemma 7 *If $2k \geq n+1$ and u is in $\mathcal{S}(R^n)$ then $S_k u$ is in $L_k^2(R^n)$ and there is a constant C, independent of u, so that*

$$\|S_k u\|_{2,k} \leq C\|u\|_{2,k} \tag{39}$$

Proof Recall that

$$\widehat{S_k u}(\xi) = U(\xi)\hat{L}_k(\xi) \tag{40}$$

where

$$U(\xi) = \sum_{\mathbf{j}\in Z^n} \hat{u}(\xi - 2\pi\mathbf{j}) = (2\pi)^{-n/2} \sum_{\mathbf{j}\in Z^n} u(\mathbf{j})e^{-i<\mathbf{j},\xi>}.$$

Observe that by virtue of (36), (40), and Plancherel's formula (39) is equivalent to

$$\int_{R^n} |\xi|^{2k}|U(\xi)\hat{L}_k(\xi)|^2 d\xi \leq C\int_{R^n} |\xi|^{2k}|\hat{u}(\xi)|^2 d\xi. \tag{41}$$

That $S_k u$ is in $L_k^2(R^n)$ follows from the readily transparent fact that the right hand side of (41) is finite.

The remainder of this proof is devoted to demonstrating (41). This demonstration involves verifying the two inequalities

$$\int_{R^n} |\xi|^{2k}|U(\xi)\hat{L}_k(\xi)|^2 d\xi \leq A\int_{Q^n} |\xi|^{2k}|U(\xi)|^2 d\xi \tag{42}$$

and

$$\int_{Q^n} |\xi|^{2k}|U(\xi)|^2 d\xi \leq B\int_{R^n} |\xi|^{2k}|\hat{u}(\xi)|^2 d\xi \tag{43}$$

where A and B are constants independent of u and Q^n is the cube

$$Q^n = \{\xi \,:\, -\pi < \xi_j \leq \pi,\ j = 1,\ldots,n\}.$$

It should be clear that (42) and (43) imply (41).

To see (42) recall that $|\xi|^{2k}|\hat{L}_k(\xi)|^2 = \hat{\Phi}_k(\xi)\hat{L}_k(\xi)$, where $\hat{\Phi}_k$ is the periodic function defined by (3); if necessary, consult [1] for more details concerning this function. Let $Q_{\mathbf{j}}^n = 2\pi\mathbf{j} + Q^n$ and write

$$\begin{aligned} \int_{R^n} |\xi|^{2k}|U(\xi)\hat{L}_k(\xi)|^2 d\xi &= \sum_{\mathbf{j}\in Z^n}\int_{Q_{\mathbf{j}}^n} |U(\xi)|^2\hat{\Phi}_k(\xi)\hat{L}_k(\xi)d\xi \\ &= \sum_{\mathbf{j}\in Z^n}\int_{Q^n} |U(\xi)|^2\hat{\Phi}_k(\xi)\hat{L}_k(\xi+2\pi\mathbf{j})d\xi \end{aligned} \tag{44}$$

where the last equality follows from a change of variable of integration and the periodicity of U and $\hat{\Phi}_k$. Now let $a_{\mathbf{j}}$ be the maximum of $\hat{L}_k(\xi + 2\pi\mathbf{j})$ for ξ in Q^n and note that $0 \leq \hat{\Phi}_k(\xi) = |\xi|^{2k}\hat{L}_k(\xi) \leq a_0|\xi|^{2k}$ if ξ is in Q^n, and recall that $a_{\mathbf{j}}$ decays exponentially. Thus we may write

(45)

$$\sum_{\mathbf{j}\in Z^n}\int_{Q^n} |U(\xi)|^2\hat{\Phi}_k(x)\hat{L}_k(\xi+2\pi\mathbf{j})dx \leq a_0\left(\sum_{\mathbf{j}\in Z^n} a_{\mathbf{j}}\right)\int_{Q^n} |\xi|^{2k}|U(\xi)|^2 dx.$$

Formula (44) and inequality (45) imply (42) with $A = a_0\left(\sum_{\mathbf{j}\in Z^n} a_{\mathbf{j}}\right)$.

To see (43) observe that for ξ in Q^n we may write

$$
\begin{aligned}
|\xi|^{2k}|U(\xi)|^2 &\leq \left(|\xi|^k \sum_{\mathbf{j}\in Z^n} |\hat{u}(\xi - 2\pi\mathbf{j})|\right)^2 \\
&\leq \left(\sum_{\mathbf{j}\in Z^n} b_{\mathbf{j}}|\xi - 2\pi\mathbf{j}|^k |\hat{u}(\xi - 2\pi\mathbf{j})|\right)^2 \\
&\leq \sum_{\mathbf{i}\in Z^n}\sum_{\mathbf{j}\in Z^n} b_{\mathbf{i}}|\xi - 2\pi\mathbf{i}|^k|\hat{u}(\xi - 2\pi\mathbf{i})| b_{\mathbf{j}}|\xi - 2\pi\mathbf{j}|^k|\hat{u}(\xi - 2\pi\mathbf{j})|
\end{aligned}
$$

where $b_0 = 1$ and otherwise $b_{\mathbf{j}}$ is equal to the maximum of $|\xi - 2\pi\mathbf{j}|^{-k}$ over ξ in Q^n. Integrating the last expression involving U over Q^n and observing that

$$\int_{Q^n} |\xi - 2\pi\mathbf{i}|^k|\hat{u}(\xi - 2\pi\mathbf{i})||\xi - 2\pi\mathbf{j}|^k|\hat{u}(\xi - 2\pi\mathbf{j})|d\xi \leq V_{\mathbf{i}}V_{\mathbf{j}}$$

where

$$V_{\mathbf{j}} = \left(\int_{Q^n} |\xi - 2\pi\mathbf{j}|^{2k}|\hat{u}(\xi - 2\pi\mathbf{j})|^2 d\xi\right)^{1/2}$$

allows us to write

$$\int_{Q^n} |\xi|^{2k}|U(\xi)|^2 d\xi \leq \sum_{\mathbf{i}\in Z^n}\sum_{\mathbf{j}\in Z^n} b_{\mathbf{i}}b_{\mathbf{j}}V_{\mathbf{i}}V_{\mathbf{j}} = \left(\sum_{\mathbf{j}\in Z^n} b_{\mathbf{j}}V_{\mathbf{j}}\right)^2. \tag{46}$$

Note that $2k > n$ implies that the sum $\sum_{\mathbf{j}\in Z^n} b_{\mathbf{j}}^2$ is finite, thus by virtue of (46) and Schwarz's inequality we have

$$\int_{Q^n} |\xi|^{2k}|U(\xi)|^2 d\xi \leq \left(\sum_{\mathbf{j}\in Z^n} b_{\mathbf{j}}^2\right)\left(\sum_{\mathbf{j}\in Z^n} V_{\mathbf{j}}^2\right). \tag{47}$$

Since

$$\sum_{\mathbf{j}\in Z^n} V_{\mathbf{j}}^2 = \int_{R^n} |\xi|^{2k}|\hat{u}(\xi)|^2 d\xi,$$

(47) implies (43) with $B = \sum b_{\mathbf{j}}^2$. ■

Proposition 8 *Suppose $2k \geq n+1$ and u is in $L_k^2(R^n)$. Then $S_k u$ is $L_k^2(R^n)$ and there is a constant C independent of u so that*

$$||S_k u||_{2,k} \leq C||u||_{2,k} \tag{48}$$

Proof By Lemma 7 S_k maps the dense subspace $\mathcal{S}(R^n)$ of $L_k^2(R^n)$ continuously into $L_k^2(R^n)$. Let $\tilde{S}_k$ denote the continuous extension of S_k onto all of $L_k^2(R^n)$. The proposition will follow if we can show that

$$S_k u = \tilde{S}_k u \tag{49}$$

in $L_k^2(R^n)$ for all u in $L_k^2(R^n)$.

To see (49) let P and Q be a pair of operators whose existence is guaranteed by Proposition 2. Now let u be any element of $L_k^2(R^n)$ and let $\{u_m\}$ be a sequence in $\mathcal{S}(R^n)$ converging to u in $L_k^2(R^n)$. By virtue of Lemma 6 the sequences $\{QS_k u_m\}$ and $\{S_k Q u_m\}$ converge to $Q\tilde{S}_k u$ and $S_k Q u$ respectively, uniformly on compact subsets of R^n. Since $\|QS_k u_m - S_k Q u_m\|_{2,k} = 0$ it follows that $Q\tilde{S}_k u - S_k Q u = p$, where p is a polynomial in $\pi_{k-1}(R^n)$. Hence

$$\tilde{S}_k u - S_k u = q, \tag{50}$$

where q is a polynomial in $\pi_{k-1}(R^n)$. From (50) it follows that $S_k u$ is in $L_k^2(R^n)$ and satisfies (49). ∎

Theorem 9 *The mapping $u \to S_k u$ is an orthogonal projection of $L_k^2(R^n)$ onto $SH_k(R^n) \cap L_k^2(R^n)$.*

Proof In view of Proposition 8 it suffices to show that S_k is idempotent and self-adjoint.

Since $S_k u$ is a k-harmonic cardinal spline for any u in $L_k^2(R^n)$ and, by virtue of the uniqueness of cardinal interpolation, $S_k u = u$ for all u in $SH_k(R^n) \cap L_k^2(R^n)$ it follows that $S_k(S_k u) = S_k u$ for all u in $L_k^2(R^n)$ and hence S_k is idempotent.

To see that S_k is self-adjoint let u and v be any elements of $\mathcal{S}(R^n)$ and, as in the proof of Lemma 7, let U and V be the periodizations of $\hat{u}$ and $\hat{v}$, namely,

$$U(\xi) = \sum_{\mathrm{j} \in Z^n} \hat{u}(\xi - 2\pi \mathrm{j})$$

and a similar formula for V. Recall that $\widehat{S_k u} = U \hat{L}_k$ and $|\xi|^{2k} \hat{L}_k(\xi) = \hat{\Phi}_k(\xi)$, use Plancherel's formula and the fact that U, V, and $\hat{\Phi}_k$ are periodic, and write

$$\begin{aligned}
\langle S_k u, v \rangle_k &= \int_{R^n} |\xi|^{2k} U(\xi) \hat{L}_k(\xi) \overline{\hat{v}(\xi)} d\xi = \int_{R^n} U(\xi) \hat{\Phi}_k(\xi) \overline{\hat{v}(\xi)} d\xi \\
&= \int_{Q^n} U(\xi) \overline{\hat{\Phi}_k(\xi) V(\xi)} d\xi = \int_{R^n} \hat{u}(\xi) \overline{\hat{\Phi}_k(\xi) V(\xi)} d\xi \\
&= \int_{R^n} |\xi|^{2k} \hat{u}(\xi) \overline{V(\xi) \hat{L}_k(\xi)} d\xi = \langle u, S_k v \rangle_k,
\end{aligned}$$

where Q^n is the cube $\{\xi \,:\, \pi < \xi_j \leq \pi,\ j = 1, \ldots, n\}$. Hence

$$\langle S_k u, v \rangle_k = \langle u, S_k v \rangle_k \tag{51}$$

holds for u and v in $\mathcal{S}(R^n)$. Since $\mathcal{S}(R^n)$ is dense in $L_k^2(R^n)$ and S_k is continuous it follows that (51) holds for all u and v in $L_k^2(R^n)$ and thus S_k is self-adjoint. ∎

Theorem 9 together with elementary facts concerning orthogonal projections on Hilbert spaces imply some interesting facts concerning k-harmonic splines in this class. We list several transparent corollaries.

Proposition 10 *If $2k \geq n + 1$ then $SH_k(R^n) \cap L_k^2(R^n)$ is a closed subspace of $L_k^2(R^n)$.*

Proposition 11 *If* $2k \geq n+1$ *then the following holds for all* u *in* $L_k^2(R^n)$:

$$\|u\|_{2,k} = \|u - S_k u\|_{2,k} + \|S_k u\|_{2,k} \tag{52}$$

and thus

$$\|S_k u\|_{2,k} \leq \|u\|_{2,k}. \tag{53}$$

Suppose u is any element in $L_k^2(R^n)$ and consider the sequence of values $\{u(j)\}$, j in Z^n. We define M_u to be that subset of $L_k^2(R^n)$ consisting of those elements v such that $v(j) = u(j)$ for all j in Z^n. Clearly M_u is an affine subspace of $L_k^2(R^n)$. In view of (53) it is not difficult to see the following concerning M_u and S_k.

Proposition 12 *If* $2k \geq n+1$ *then there exists a unique element* w *in* M_u *such that*

$$\|w\|_{2,k} = \min_{v \in M_u} \|v\|_{2,k},$$

and $w = S_k u$. *In other words,* $S_k u$ *is the unique element in* M_u *of minimal* $L_k^2(R^n)$ *norm.*

4 Cardinal interpolation in $L_k^2(R^n)$ and $\ell_k^2(Z^n)$.

As mentioned in the introduction, in this section we present necessary and sufficient conditions on a sequence $\{v_j\}$, j in Z^n, which allow it to be interpolated by the elements of $L_k^2(R^n)$. In addition the interpolating element of minimal $L_k^2(R^n)$ norm is characterized.

First recall the definitions of the class $\mathcal{F}$ of finite difference operators given earlier. We say that T in $\mathcal{F}$ is of order k if $\hat{T}(\xi) = O(|\xi|^k)$ but not $o(|\xi|^k)$ as ξ goes to 0.

Proposition 13 *Suppose* $2k \geq n+1$, u *is in* $L_k^2(R^n)$, *and* T *is any finite difference operator of order* $\geq k$. *Then*

$$\sum_{j \in Z^n} |Tu(j)|^2 \leq C\|u\|_{2,k}^2 \tag{54}$$

where C *is a constant independent of* u.

Proof If u is in $\mathcal{S}(R^n)$, let

$$U(\xi) = (2\pi)^{-n/2} \sum_{j \in Z^n} u(j) e^{-i(j,\xi)}$$

and recall the definition of $\Phi_k(\xi)$. By virtue of Parseval's identity and the fact that $|\hat{T}(\xi)|^2/\hat{\Phi}_k(\xi)$ is bounded we may write

$$\begin{aligned}
\sum_{j \in Z^n} |Tu(j)|^2 &= \int_{Q^n} |\hat{T}(\xi)U(\xi)|^2 d\xi \\
&= \int_{Q^n} |U(\xi)|^2 |\hat{T}(\xi)|^2 |\hat{\Phi}_k(\xi)|^2 |\hat{\Phi}_k(\xi)|^{-2} d\xi \\
&\leq C \int_{Q^n} |U(\xi)|^2 |\hat{\Phi}_k(\xi)|^2 |\hat{\Phi}_k(\xi)|^{-1} d\xi
\end{aligned}$$

$$
\begin{aligned}
&= C\int_{R^n} |U(\xi)\hat{\Phi}_k(\xi)|^2|\xi|^{-2k}d\xi \\
&= C\int_{R^n} |U(\xi)\hat{\Phi}_k(\xi)|\xi|^{-2k}|^2|\xi|^{2k}d\xi \\
&= C\|S_k u\|_{2,k}\ .
\end{aligned}
$$

Since $\|S_k u\|_{2,k} \leq \|u\|_{2,k}$ we may conclude that (54) holds whenever u is in $\mathcal{S}'(R^n)$. To see that (54) holds for all u in $L^2_k(R^n)$ let P and Q be the operators whose existence is guaranteed by Proposition 2, recall that $Tp = 0$ for all polynomials in $\pi_{k-1}(R^n)$, and observe that

$$TQu = Tu \tag{55}$$

for all u in $L^2_k(R^n)$. Choose any element u in $L^2_k(R^n)$, and let $\{u_m\}$ be a sequence in $\mathcal{S}(R^n)$ converging to u in $L^2_k(R^n)$. By virtue of Lemma 6 $\{Qu_m\}$ converges to Qu uniformly on compact subsets of R^n and hence in view of (55) $\{Tu_m(\mathbf{j})\}$ converges to $Tu(\mathbf{j})$ for all $\mathbf{j}$ in Z^n. This last observation together with the fact that (54) holds for each u_m implies that

$$\sum |Tu(\mathbf{j})|^2 \leq C\|u\|^2_{2,k}$$

where the sum is taken over any finite subset of Z^n. The last inequality of course implies the desired result. ■

We say that a finite collection T_m , $m = 1, \ldots, N$, of finite differences satisfies condition τ_k if

$$\hat{\Phi}_k(\xi)\left(\sum_{m=1}^{N} |\hat{T}_m(\xi)|^2\right)^{-1}$$

is in $L^\infty(R^n)$. Note that the T^ν 's used in the definition of $\ell^2_k(Z^n)$ enjoy this property. Also note that for every positive integer k there is a finite collection in $\mathcal{F}$ which satisfies condition τ_k.

Proposition 14 *Suppose $2k \geq n+1$, u is a sequence in $\ell^2_k(Z^n)$, and f_u is the unique k-harmonic cardinal spline which interpolates u. Then f_u is in $L^2_k(R^n)$ and*

$$\|f_u\|_{2,k} \leq C\|u\|_{2,k} \tag{56}$$

where C is a constant independent of u.

Proof Suppose u is in $\mathcal{Y}^{-\infty}$, then

$$f_u = \sum_{\mathbf{j}\in Z^n} u_{\mathbf{j}} L_k(x - \mathbf{j}),$$

the function

$$U(\xi) = \sum_{\mathbf{j}\in Z^n} u_{\mathbf{j}} e^{-i(\mathbf{j},\xi)}$$

is well defined, and $\widehat{f_u} = U\hat{L}_k$. Applying Plancherel's formula and the fact that the collection $\{T^\nu : |\nu| = k\}$ satisfies condition τ_k we may write

$$\|f_u\|^2_{2,k} = \int_{R^n} |U(\xi)\hat{L}_k(\xi)|^2|\xi|^{2k}d\xi$$

$$
\begin{aligned}
&= \int_{R^n} |U(\xi)|^2 \hat{\Phi}_k^2(\xi)\hat{L}_k(\xi)d\xi \\
&= \int_{Q^n} |U(\xi)|^2 \hat{\Phi}_k(\xi)d\xi \\
&= \int_{Q^n} |U(\xi)|^2 \left\{\sum_{\nu=k} |\widehat{T^\nu}(\xi)|^2\right\} \left\{\frac{\hat{\Phi}_k(\xi)}{\sum_{|\nu|=k} |\widehat{T^\nu}(\xi)|^2}\right\} d\xi \\
&\leq C\int_{Q^n} \sum_{|\nu|=k} |\widehat{T^\nu}(\xi)U(\xi)|^2 d\xi \\
&= C\|u\|_{2,k}^2 .
\end{aligned}
$$

Thus (56) holds for u in $\mathcal{Y}^{-\infty}$. In view of Propositions 3 and 4 the desired result follows from a density argument similar to that used in the proof of Proposition 13. ■

Corollary 1 *The mapping*

$$\{u_{\mathrm{j}}\} \longrightarrow f_u = \sum_{\mathrm{j}\in Z^n} u_{\mathrm{j}} L_k(x-\mathrm{j})$$

is an isomorphism between $\ell_k^2(Z^n)$ *and* $L_k^2(R^n)\cap SH_k(R^n)$ *such that*

$$c\|u\|_{2,k} \leq \|f_u\|_{2,k} \leq C\|u\|_{2,k}$$

where c *and* C *are positive constants independent of* u.

We conclude this paper by summarizing the contents of the above two propositions in the following theorems.

Theorem 15 *Given a sequence* $u = \{u_{\mathrm{j}}\}$, *there is an element in* $L_k^2(R^n)$ *which interpolates it if and only if* u *is in* $\ell_k^2(Z^n)$.

Combining the results in this section together with Propositions 11 and 12 easily produces the following conclusion.

Theorem 16 *Suppose* u *is in* $\ell_k^2(Z^n)$. *Then there is a unique k-harmonic spline* f_u *in* $L_k^2(R^n)\cap SH_k(R^n)$ *which interpolates* u. *This interpolant* f_u *has the property that*

$$\|f_u\|_{2,k} < \|g\|_{2,k}$$

for any g *in* $L_k^2(R^n)$ *which interpolates* u, *unless* $g(x) = f_u(x)$ *for all* x *in* R^n. *In other words,* f_u *is the unique element of minimal* $L_k^2(R^n)$ *norm which interpolates* u.

References

[1] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines, *Journal of Approximation Theory*, to appear.

[2] I. J. Schoenberg, *Cardinal Spline Interpolation*, CBMS Vol. 12, SIAM, Philadelphia, 1973.

Error bounds for multiquadric interpolation

W. R. Madych and S. A. Nelson

Abstract

A class of multivariate scattered data interpolation methods which includes the so-called multiquadrics is considered. Pointwise error bounds are given in terms of several parameters including a parameter d which, roughly speaking, measures the spacing of the points at which interpolation occurs. These bounds can be of arbitrarily high order in d.

§1. Introduction

To avoid technical complications which are not germaine to the basic ideas we restrict the major part of this report to the following simple setup.

Let h be the function defined for x in R^n, $n \geq 1$, by the formula

$$h(x) = -\sqrt{1 + |x|^2} \tag{1}$$

where $|x|$ is the Euclidean norm of x. Given data (x_j, f_j), $j = 1, \dots, N$, where $X = \{x_1, \dots, x_N\}$ is a subset of points in R^n and the f_j's are real or complex numbers, consider the function s defined by

$$s(x) = c_o + \sum_{j=1}^{n} c_j h(x - x_j) \tag{2}$$

where the c_j's are chosen so that

$$\sum_{j=1}^{n} c_j = 0 \text{ and } c_o + \sum_{j=1}^{N} c_j h(x_i - x_j) = f_i, \quad i = 1, \dots, N. \tag{3}$$

It is well known that the system of equations (3) has a unique solution and thus the interpolant $s(x)$ is well defined. Recall that the system

$$\sum_{j=1}^{N} a_j h(x_i - x_j) = f_i, \quad i = 1, \dots, N, \tag{4}$$

Approximation Theory VI
C. K. Chui, L. L. Schumaker and J. D. Ward (eds.)

ISBN 0-12-17585-6:

can also be solved uniquely for the coefficients a_j. However the resulting interpolant

$$\sum_{j=1}^{N} a_j h(x - x_j)$$

is not necessarily in the class C_h. This class, which is defined below, plays an important role in our estimates.

The method of interpolation (3) or (4) is often referred to as the multiquadric method. For more background and further references see the survey article by Nira Dyn in these proceedings.

Recall that the generalized Fourier transform of h is given by

$$\hat{h}(\xi) = \frac{c_n K_\lambda(|\xi|)}{|\xi|^\lambda} \tag{5}$$

where c_n is a constant which depends on n, K_λ is a Bessel function of the third kind of order λ, and $\lambda = (n+1)/2$. For example, in the case $n = 2$ we may write

$$\hat{h}(\xi) = \frac{(1 + |\xi|)e^{-|\xi|}}{(2\pi)^2 |\xi|^3}.$$

Note that in the general case $\hat{h}$ has exponential decay at infinity and $\hat{h}(\xi) = O(|\xi|^{-n-1})$ at the origin.

Set $r(\xi) = 1/\{|\xi|^2 \hat{h}(\xi)\}$ and let C_h be the class of those tempered distributions f whose first order derivatives have Fourier transforms which are square integrable with respect to the weight function $r(\xi)$ over R^n. Define $\|f\|_h$ by

$$\|f\|_h^2 = \sum_{j=1}^{n} \int_{R^n} |\widehat{D_j f}(\xi)|^2 r(\xi) d\xi.$$

Note that $\|f\|_h^2$ is a positive semidefinite quadratic form on C_h and thus $\|f\|_h$ may be regarded as a seminorm on C_h. It is not difficult to see that C_h is essentially contained in the class of infinitely differentiable functions on R^n; the definition implies that derivatives of arbitrarily high order are square integrable. Another important property of C_h is the fact that the data $\{f_j\}$ can be regarded as the pointwise restriction of an element of C_h to X.

As suggested by the definition there is an intimate connection between the class C_h and the interpolant $s(x)$. We summarize some of the properties which follow from this relationship.

i. The interpolant s is in C_h.

ii. If g is in C_h and $g(x_j) = f_j$, $j = 1, \ldots, N$, then

$$\|g\|_h^2 = \|g - s\|_h^2 + \|s\|_h^2.$$

iii. The function s is the element of minimal C_h norm which interpolates the data (x_j, f_j), $j = 1, \ldots, N$. In other words

$$\|s\|_h = \min\{\|g\|_h \; : \; g \in C_h \text{ and } g(x_j) = f_j, \; j = 1, \ldots, N\}.$$

§2. Error bounds

Before stating the basic results we recall some technical definitions.

In what follows the discrete set X is always assumed to be a subset of an open set Ω. Furthermore, there are constants M and ϵ_0 such that for every ϵ, $0 < \epsilon < \epsilon_0$,

$$\Omega \subset \cup\{B(x, \epsilon M) \, : \, x \in \Omega_\epsilon\} \tag{6}$$

where $\Omega_\epsilon = \{x \in R^n \, : \, B(x,\epsilon) \subset \Omega\}$ and $B(x,\epsilon) = \{y \in R^n \, : \, |y - x| < \epsilon\}$. It is not difficult to see that (6) is satisfied if Ω is an ellipsoid or parallelpiped in R^n; more generally, (6) is satisfied if Ω satisfies an interior cone condition. Finally we define the parameter d by

$$d = \sup_{y\in\Omega} \inf_{x\in X} |y - x|.$$

Note that d depends on both Ω and X; it is a standard measure of how well Ω is covered by X.

Theorem. *Suppose f is in C_h and $f(x_j) = f_j$, $j = 1, \ldots, N$. Given an integer $k \geq 1$ there are constants $C = C(k, \Omega)$ and $d_0 = d(k, \Omega)$ such that if $d < d_0$ and the multi-index ν satisfies $|\nu| \leq k$ then*

$$\|D^\nu(f - s)\|_{L^\infty(\Omega)} \leq C d^{k-|\nu|} \|f\|_h.$$

Of course the norm on the left hand side of the above inequality is just the usual L^∞ norm taken over Ω. We also remind the reader that, whereas $|x|$ normally denotes the Euclidean norm of x, in the case of multi-indicies $|\nu|$ denotes the sum of the components of ν.

The argument used to obtain this theorem can be outlined as follows.

First, by virtue of the properties *(i)-(iii)* listed above, there is no loss of generality when attention is restricted to the case $s = 0$. In this case the proof can be broken down into four steps.

If σ_x is a compactly supported Borel measure on R^n such that $p(x) = \int p(y) d\sigma_x(y)$ for all polynomials p of degree less than k then

$$|f(x) - \int f(y) d\sigma_x(y)| \leq c \|f\|_h \int |y - x|^k d|\sigma_x|(y), \tag{7}$$

where c depends only on h, k, and n. Use reasoning similar to that used in [1] and apply (7) locally in Ω with σ's supported in a (small ball around x)$\cap X$ to conclude that

$$|f(x)| \leq C d^k \|f\|_h \tag{8}$$

for all x in Ω. Next, if f is in C_h then

$$\|D^\nu f\|_{L^\infty} \leq c \|f\|_h \tag{9}$$

for $|\nu| \geq 1$, where the L^∞ norm is over all of R^n and c is a constant which depends only on h, n, and ν. Finally, 'interpolate' between (8) and (9) to obtain the desired result.

Details can be found in [2]

§3. Extensions and limitations

The error bound given above holds for methods of interpolation of the form (2) and (3) where h of (1) is replaced by a more general function which is *conditionally positive definite of order* m. Of course the space $\mathcal{C}_h$ must change accordingly and certain limitations on k may apply.

To be more specific, recall that such an h uniquely determines a non-negative Radon measure μ on $R^n \setminus \{0\}$ with $\mu = \hat{h}$ on $R^n \setminus \{0\}$. If for some ℓ, $\ell \geq m$, this measure satisfies

$$\int_{|\xi|>1} |\xi|^{2\ell} d\mu(\xi) < \infty, \tag{10}$$

then the theorem stated above holds with the restriction $m \leq k \leq \ell$.

For example, if $h(x) = (1+|x|^2)^{\alpha}$ where α is any fixed real number which is not a non-negative integer then h or $-h$ is conditionally positive definite of order m where m is any integer greater than α. In this case the Fourier transform of h is smooth away from the origin and decays exponentially at infinity. Thus (7) holds for every integer ℓ and the appropriate version of the theorem holds with $\mathcal{C}_h$ defined accordingly whenever the integer k satisfies $\alpha \leq k < \infty$.

The discrete set X need not be finite and Ω need not be bounded. An important requirement however is that the numbers $\{f_j\}$ be the values $\{f(x_j)\}$ of some function f in $\mathcal{C}_h$. In the infinite case this requirement may not be so easy to verify. Nevertheless interesting applications of this result can be had in the case of infinite X; for example, in various so-called cardinal methods of interpolation.

References

1. Duchon, J., Sur l'erreur d'interpolation des fonctions de plusiers variables par les D^m-splines, *RAIRO Analyse Numerique* 12 (1978), 325-334.
2. Madych, W. R., and S. A. Nelson, Multivariate Interpolation and conditionally positive definite functions II, preprint.

Both authors were partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

On the Correctness of
the Problem of Inverting the Finite Hilbert Transform
in Certain Aeroelastic Models

W. R. Madych*

Abstract

We indicate methods of ensuring that problem in the title is correctly posed in the L^p sense whenever the derivative of the circulation function satisfies certain mild conditions.

1 Introduction

In the theory of aeroelastic control systems it is required to solve

$$f(x) = \int_{-1}^{1} \frac{\gamma(y)}{y-x}\, dy \tag{1}$$

for $\gamma(y)$, $-1 < y < 1$, in terms of $f(x)$, $-1 < x < 1$. The function f is assumed to be of the form

$$f(x) = w(x) + g(x)\ , \tag{2}$$

where

$$g(x) = \int_{0}^{\infty} \frac{G(s)}{1-x+s}\, ds. \tag{3}$$

Here w and G are constant multiples of the so-called downwash function and the derivative of the circulation function respectively.

Formula (1) is often referred to as the finite Hilbert transform of γ, see [3]. For more detail concerning this model of aeroelasticity see [1] and the references cited there. In particular, $f = f^t$ and $G = \dot{\Gamma}_t$ in the notation of [1].

In order to guarantee the correctness of the problem of solving (1) via the methods in [3], it is necessary to assume that f is in some $L^p(-1,1)$ class[1].

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

[1]If J is an interval and p is a positive number then $L^p(J)$ is the class of those Lebesgue measurable functions for which $\int_J |G(x)|^p dx$ is finite. When $p = \infty$, $L^\infty(J)$ is the class of essentially bounded functions on J.

In view of this and the fact that w can usually be taken to be in any class $L^p(-1,1)$, it is important to obtain a fairly general answer to the following question: What conditions on G ensure that g is in $L^p(-1,1)$?

It is the purpose of this note to point out certain natural methods for obtaining such conditions. Propositions 2 and 3 below contain several typical results.

2 Discussion

Observe that the integral defining $g(x)$ is a smooth function of x for $x < 1$ whenever G is locally integrable and satisfies a mild condition at infinity. In particular, it is clear that $g(x)$ is infinitely differentiable for $x < 1$ if

$$\int_0^\infty \frac{|G(s)|}{1+s}\, ds < \infty. \tag{4}$$

In fact, if $0 < x < 1$, we may write

$$|g(x)| \leq (1-x)^{-1} \int_0^\infty \frac{|G(s)|}{1+s}\, ds$$

and conclude that the only questionable behavior of g occurs in arbitrarily small neighborhoods of $x = 1$ whenever f satisfies (4).

It should be noted that condition (4) is quite mild and general. For example, if $G = G_1 + G_2$ where G_1 is in $L^1(0,\infty)$ and G_2 is in $L^p(0,\infty)$ for some p, $p < \infty$, then G satisfies (4).

To understand how G influences the behavior of g in neighborhoods of $x = 1$, express the integral defining g as a sum, $\int_0^\epsilon + \int_\epsilon^\infty$, where ϵ is any positive number ≤ 1. Since

$$\left|\int_\epsilon^\infty \frac{G(s)}{1-x+s}\, ds\right| \leq \frac{2}{\epsilon}\int_0^\infty \frac{|G(s)|}{1+s}\, ds\,, \tag{5}$$

it should be clear that the behavior of g at $x = 1$ is determined by the behavior of G at the origin. Indeed, if $s^{-\alpha}G(s)$ is in $L^p(0,\epsilon)$ for some value of p, $1 \leq p \leq \infty$, then by virtue of Holder's inequality we may write

$$\left|\int_0^\epsilon \frac{G(s)}{1-x+s}\, ds\right| \leq I(x,p,\alpha)\left[\int_0^\epsilon \left(s^{-\alpha}G(s)\right)^p ds\right]^{1/p}, \tag{6}$$

where

$$I(x,p,\alpha) = \left[\int_0^\epsilon \left(\frac{s^\alpha}{1-x+s}\right)^{p/(p-1)} ds\right]^{1-1/p}.$$

Now, using a change of variable, I may be expressed as

$$I(x,p,\alpha) = (1-x)^{\alpha-1/p}\left[\int_0^{\epsilon/(1-x)} \left(\frac{s^\alpha}{1+s}\right)^{p/(p-1)} ds\right]^{1-1/p},$$

from which we may easily estimate its size.

We summarize these observations as follows:

Proposition 1 *Suppose g is related to G via (3), G satisfies condition (4), and $s^{-\alpha}G(s)$ is in $L^p(0,\epsilon)$ for some positive ϵ, some α, $\alpha \geq 0$, and some p, $1 \leq p \leq \infty$. Then for $-1 < x < 1$*

$$|g(x)| \leq \begin{cases} C(1-x)^{\alpha - 1/p} & \text{if } \alpha < 1/p \\ C(1+\log(1-x)) & \text{if } \alpha = 1/p \\ C & \text{if } \alpha > 1/p \end{cases} \tag{7}$$

where C is independent of x.

A result concerning the L^p class of g follows as an immediate corollary.

Proposition 2 *Suppose G and g satisfy the hypothesis of Proposition 1. If $\alpha \geq 1/p$ then g is in $L^q(-1,1)$ for all positive q. If $\alpha < 1/p$ then g is in $L^q(-1,1)$ for all positive q which satisfy $q < p/(1-\alpha p)$.*

By using a slightly more delicate argument, the inequality $q < p/(1-\alpha p)$ in the second half of the above proposition can be tightened to $q \leq p/(1-\alpha p)$ in the case $1 < p < \infty$. To see this, use the fact that if $\alpha \geq 0$ and $x < 1$ then

$$\left(\frac{s}{1-x+s}\right)^{\alpha} \leq 1$$

to observe that

$$\left|\int_0^{\epsilon} \frac{G(s)}{1-x+s}\, ds\right| \leq I_\alpha G_\epsilon^\alpha(x-1) \tag{8}$$

where

$$I_0\phi(y) = \int_{-\infty}^{\infty} \frac{\phi(s)}{s-y}\, ds$$

is the classical Hilbert transform of ϕ and, when $\alpha > 0$,

$$I_\alpha\phi(y) = \int_{-\infty}^{\infty} \frac{\phi(s)}{|s-y|^{1-\alpha}}\, ds$$

is the fractional integral or Riesz potential of ϕ. Here

$$G_\epsilon^\alpha(s) = \begin{cases} |s^{-\alpha}G(s)| & \text{if } 0 < s < \epsilon \\ 0 & \text{if } s < 0 \text{ or } s > \epsilon\,. \end{cases}$$

The mapping properties of the transformation $\phi \rightarrow I_\alpha\phi$ are well known and, in view of (8), can be used to make conclusions concerning the behavior of g. For instance, if $1 < p < \infty$, $0 \leq \alpha < 1/p$, and ϕ is in $L^p(-\infty,\infty)$ then $I_\alpha\phi$ is in $L^q(-\infty,\infty)$ where $q = p/(1-\alpha p)$; see [2]. This together with (5), (8), and Proposition 1 allows us to conclude the following:

Proposition 3 *Suppose G and g satisfy the hypothesis of Proposition 1 with the restriction that $1 < p < \infty$ and $0 \leq \alpha < 1/p$. Then g is in $L^q(-1,1)$ for all positive q such which satisfy $q \leq p/(1-\alpha p)$.*

Another method of estimating the left hand side of (8) involves writing $\int_0^\epsilon$ as $\int_0^{(1-x)} + \int_{(1-x)}^\epsilon$ when $1 - x < \epsilon$, using the change of variables $y = 1 - x$, and applying variants of Hardy's inequality, see [2] page 245, to each of the resulting integrals. This leads to generalizations of Proposition 1 for certain values of the parameters α and p.

For the sake of completeness we mention that, using similar methods involving fractional integrals, it is possible to obtain results concerning the behavior of g in neighborhoods of 1 for other values of the parameters α and p. Since such estimates require the introduction of certain technical machinery and the conclusions do not involve L^p, we will not pursue the details here.

It may be worth noting that in the notation of [1] $G(s)$ is equal to $\psi(t-s)$ for $t < s$; it is equal to another expression when $s < t$. Furthermore ψ is assumed to be in $L^1(-\infty, 0)$. In this case our observations imply that if $t > 0$ then ψ does not affect the L^p class of f^t. This should be compared with the conclusions in [1].

References

[1] J. A. Burns, E. M. Cliff, T. L. Herdman and J. Turi, On integral transforms appearing in the derivative of the equations of an aeroelastic system, in *Nonlinear Analysis and Applications*, V.Lakshmikantham, ed. Lecture Notes in Pure and Applied Mathematics, Volume 109, Marcel Dekker, Inc., New York, (1987) 89-98.

[2] G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, 2nd ed., Cambridge, 1952.

[3] F. G. Tricomi, *Integral equations*, Interscience Publishers, Inc. New York, 1957.

Cardinal Interpolation with Polyharmonic Splines

W. R. Madych*

1 Introduction

If k is a positive integer an n-variate *k-harmonic cardinal spline* is a tempered distribution f on R^n such that $\Delta^k f$ is a measure supported on the integer lattice Z^n in R^n. Symbolically

$$\Delta^k f(x) = \sum_{\mathbf{j}\in Z^n} a_{\mathbf{j}}\delta(x - \mathbf{j}) \tag{1}$$

where Δ is the n variate Laplacian, $\Delta^k = \Delta\Delta^{k-1}$ for $k > 1$, and $\delta(x)$ is the unit Dirac measure supported at the origin. A *polyharmonic spline* is one which is k-harmonic for some k.

The most important use of these distributions involves the interpolation of data $\{u_{\mathbf{j}}\}$ defined on Z^n to all of R^n. Thus the *cardinal interpolation problem* for k-harmonic splines is the following: given a sequence $\{u_{\mathbf{j}}\}$ defined on Z^n find a k-harmonic spline f such that

$$f(\mathbf{j}) = u_{\mathbf{j}} \tag{2}$$

for all $\mathbf{j}$. Besides the basic questions of existence and uniqueness other meaningful questions concern the behavior of the spline interpolant in terms of the data and the effect of the parameter k.

In the univariate case, $n = 1$, these splines are exactly the polynomial cardinal splines of odd degree and polynomial growth studied by Schoenberg, [20].

It is the purpose of this note to describe the properties of these splines in the general n-variate case. Except for the existence of B-splines, these properties are remarkably similar to those found in the univariate case.

Some of these results were originally announced in [10] and amplified in [11] where motivation and other background is given. Here we only mention that the distributional definition of these splines which is given above is motivated by

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

the facts that it is very convenient and easily allows for the unrestrained use of Fourier transforms. Thus in a certain sense this development may be regarded as an extension of the early work of Schoenberg [19] which used Fourier analysis.

Recall that the classical univariate cardinal splines of degree $2k-1$ and polynomial growth may be viewed as linear combinations of translates of $|x|^{2k-1}$. Similarly the general k-harmonic cardinal spline f may be regarded as a linear combination of translates of the fundamental solution of Δ^k denoted by $E_k(x)$; thus $\Delta^k E_k(x) = \delta(x)$ and

$$f(x) = \sum_{\mathrm{j}\in Z^n} a_{\mathrm{j}} E_k(x - \mathrm{j}) . \tag{3}$$

Note that at this point the correspondence between (1) and (3) is merely symbolic; clearly (1) is meaningful in the tempered distribution sense for any sequence $\{a_{\mathrm{j}}\}$ of polynomial growth whereas this is not so apparent for (3).

Interpolation in terms of linear combinations of translates of a fixed function h is certainly very appealing. It arises naturally when the interpolants are solutions of certain variational problems: the general idea goes at least as far back as that of the reproducing kernel Hilbert space, for example see [1,8], but the first meaningful use of the Fourier transform in this context seems to be due to Duchon, [6]; usually in such cases h is the fundamental solution of an appropriate differential or pseudo-differential operator and the interpolant is often further modified in some way to account for various auxiliary conditions and restrictions, for *example see* [6,7,13,15,17,21] and the appropriate references therein. The classes of positive definite and conditionally positive definite functions provide a rich collection of examples of h's which can be used for such interpolation; various dilates of the Gaussian $\exp(-|x|^2)$, the so-called multiquadric $\sqrt{1+|x|^2}$, and the fundamental solution $E_k(x)$, $2k \geq n+1$, are specific examples; for for details and more examples see [2,7,13,16]. When the data is given on a lattice the univariate B-splines are perhaps the most popular example of such h's, see [20,18]; the recently developed theory of box splines seems to be an attempt to generalize this concept to the multi-variate case, see [3,5] and the pertinent articles in this volume.

The notation used here is standard, if necessary see [11] for a more detailed explanation. Here we merely remind the reader that there are several common normalizations for the Fourier transform. In this note we use

$$\hat{\varphi}(\xi) = (2\pi)^{-n/2} \int_{R^n} \varphi(x) e^{-i\langle \xi, x\rangle} dx$$

for the Fourier transform $\hat{\varphi}$ of a test function φ.

2 Basic properties

Suppose f is a k-harmonic cardinal spline. Then f is clearly analytic on $R^n \setminus Z^n$ and, in order to ensure that point evaluation on Z^n is meaningful, we assume

that f is continuous on all of R^n. In this case, if $2k < n+1$ it is well known that f must be a k-harmonic polynomial; this follows from the behavior of the corresponding fundamental solution at the origin, see [9] for details. Now, the class of k-harmonic polynomials is too exclusive to interpolate a sufficiently broad class of data $\{u_{\mathbf{j}}\}$ on Z^n. For example, it is not difficult to see that there is no k-harmonic polynomial f such that $f(0) = 1$ and $f(\mathbf{j}) = 0$ for $\mathbf{j}$ in $Z^n \setminus \{0\}$. For this reason we restrict our attention to the case $2k \geq n+1$ in what follows.

Let $SH_k(R^n)$ denote the class of k-harmonic splines on R^n and recall that $C^k(R^n)$ is the class of functions which are k times continuously differentiable on R^n.

Proposition 1 *If $2k \geq n+1$ and f is in $SH_k(R^n)$ then f is in $C^{2k-n-1}(R^n)$.*

Recall that a sequence $\{u_{\mathbf{j}}\}$ is of polynomial growth if there are constants c and p such that

$$|u_{\mathbf{j}}| \leq c(1+|j|)^p \tag{4}$$

holds for all $\mathbf{j}$ in Z^n. The class $\mathcal{Y}^\alpha$ is the collection of those sequences for which (4) holds for $p = \alpha$. Similarly a continuous function f is of polynomial growth if there are constants c and p such that

$$|f(x)| \leq c(1+|x|)^p \tag{5}$$

holds for all x in R^n. The class $SH_k^\alpha(R^n)$ is the collection of those k-harmonic splines for which (5) holds for $p = \alpha$.

Proposition 2 *If $2k \geq n+1$ and f is in $SH_k(R^n)$ then f is of polynomial growth. In other words $SH_k(R^n) = \bigcup SH_k^\alpha(R^n)$ where the union is taken over all $\alpha < \infty$.*

The last proposition implies that in the univariate case the class $SH_k(R)$ is not quite as general as the corresponding class $\mathcal{S}_{2k-1}$ considered by Schoenberg. It fails to contain the subspace of $\mathcal{S}_{2k-1}$ consisting of the null-splines, see [20, Lecture 4, Section 3].

Since the elements of $SH_k(R^n)$ are of polynomial growth it is clear that in order for cardinal interpolation problem to have a solution a necessary requirement on the data sequence is that it also be of polynomial growth. As we shall see this requirement is also sufficient. We begin by first considering the fundamental functions of interpolation.

Consider the distribution L_k which is defined by the formula for its Fourier transform:

$$\hat{L}_k(\xi) = (2\pi)^{-n/2} \frac{|\xi|^{-2k}}{\sum_{\mathbf{j}\in Z^n} |\xi - 2\pi\mathbf{j}|^{-2k}}. \tag{6}$$

If k is an integer such that $2k \geq n+1$ then $L_k(x)$ is well defined as an absolutely convergent integral.

Proposition 3 *Let L_k be defined by the formula for its Fourier transform (6), where k is an integer which satisfies $2k \geq n+1$. Then L_k has the following properties:*

(i) L_k is a k-harmonic cardinal spline.

(ii) For all $\mathbf{j}$ *in Z^n*

(7) $$L_k(\mathbf{j}) = \begin{cases} 1 & \text{if } \mathbf{j} = 0 \\ 0 & \text{if } \mathbf{j} \neq 0 \end{cases}$$

(iii) There are positive constants A and a, depending on n and k but independent of x, such that

(8) $$|L_k(x)| \leq A \exp(-a|x|)$$

for all x in R^n.

(iv) L_k has the following representations in terms of E_k:

(9) $$L_k(x) = \sum_{\mathbf{j} \in Z^n} a_{\mathbf{j}} E_k(x - \mathbf{j})$$

where the $a_{\mathbf{j}}$'s satisfy $|a_{\mathbf{j}}| \leq B \exp(-b|\mathbf{j}|)$ and the positive constants B and b depend only on n and k. The series converges absolutely and uniformly on all compact subsets of R^n.

Item *(i)* is a readily transparent consequence of the definition while item *(ii)* easily follows from the fact that for $\mathbf{j}$ in Z^n

(10) $$L_k(\mathbf{j}) = \frac{1}{2\pi} \int_{Q^n} e^{i(\mathbf{j}, \xi)} d\xi$$

where

$$Q^n = \{\xi = (\xi_1, \ldots, \xi_n) \ : \ -\pi < \xi_j \leq \pi, \ j = 1, \ldots, n\}.$$

The remaining items are consequences of the analyticity of $\hat{L}_k$.

The following theorem concerns the cardinal spline interpolation problem and the nature of its solution. It is essentially a routine consequence of Proposition 3 except for the uniqueness assertion. Details may be found in [11].

Proposition 4 *If $\{u_{\mathbf{j}}\}$, $\mathbf{j}$ in Z^n, is a sequence of polynomial growth and $2k \geq n+1$ then the following is true:*

(i) There is a unique k-harmonic spline f such that $f(\mathbf{j}) = u_{\mathbf{j}}$ for all $\mathbf{j}$.

(ii) If $u_{\mathbf{j}}$ is in $\mathcal{Y}^\alpha$ then f is in $SH_k^\alpha(R^n)$.

(iii) Every k-harmonic spline f has a unique representation in terms of translates of L_k, namely

$$f(x) = \sum_{\mathbf{j} \in Z^n} f(\mathbf{j}) L_k(x - \mathbf{j}). \tag{11}$$

The expansion (11) converges absolutely and uniformly in every compact subset of R^n.

3 Other Properties

Having given the basic existence and uniqueness result concerning the cardinal interpolation problem for k-harmonic splines we briefly address the following questions:

How is the behavior of the data sequence $\{u_{\mathbf{j}}\}$ reflected in the behavior of its k-harmonic spline interpolant?

Most of the results known in the univariate case have an appropriate analogue in the general case. For example, item *(ii)* of Proposition 4 addresses this question. Other results include appropriate analogues of [4]; namely, if $\{u_{\mathbf{j}}\}$ or certain finite differences of it are in $\ell^p(Z^n)$ the the corresponding k-harmonic spline interpolant or appropriate derivatives are in $L^p(R^n)$.

What about the variational properties of these splines?

The appropriate analogue of results cited in [20, Lecture 6, Section 1] hold in the general case. See [12] for details.

If the data sequence is fixed what is the behavior of the cardinal k-harmonic spline interpolant as $k \to \infty$.

Again, most of the results which hold in the univariate case have an appropriate analogue here. A result which seems to be new even in the univariate case is the following: Suppose f is such that the support of $\hat{f}$ is in Q^n and s_k is its cardinal k-harmonic spline interpolant, namely, $s_k(\mathbf{j}) = f(\mathbf{j})$ for $\mathbf{j}$ in Z^n. Then if $\hat{f}$ satisfies a mild condition in a neighborhood of the boundary of Q^n then

$$\lim_{k \to \infty} s_k(x) = f(x)$$

uniformly on compact subsets of R^n. For some details see [14].

Suppose $s(x)$ is the k-harmonic spline interpolant of f on the dilated lattice aZ^n. What is the degree of approximation in terms of a?

Recall that $s(x)$ reproduces any k-harmonic polynomial; in particular, it reproduces any polynomial of degree $\leq 2k-1$. This together with routine arguments involving (8) implies that $|s(x) - f(x)| = O(a^{2k})$ as $a \to 0$ whenever f is in $C^{2k}(R^n)$ with bounded derivatives of order $2k$. Similar results hold with L^p norms. Interpolation arguments imply appropriate results for less regular f's. Thus appropriate analogues of all the univariate results hold in the general case.

What about numerical implementations?

The fundamental spline $L_k(x)$ can be easily evaluated quite rapidly and accurately via the fast Fourier transform. See Figures 1 and 2. In view of this the computation of most k-harmonic spline interpolants should should pose no significant difficulties.

What about further generalizations?

There are many directions in which one can extend certain aspects of this theory. For example, by replacing $|\xi|^{-2k}$ and its periodization in formula (6) by an appropriate function ϕ which decays sufficiently fast at ∞ it is readily transparent that the corresponding analogue of L_k, call it L, will be continuous, will certainly satisfy (10), and thus will also satisfy (7). If ϕ is sufficiently regular then L will have corresponding decay properties at ∞. In view of this it is quit natural to consider interpolants of the form

$$f(x) = \sum_{\mathrm{j}\in Z^n} f(\mathrm{j})L(x-\mathrm{j})\ .$$

Typical ϕ's for which the basic properties of the above type interpolants are particularly transparent are $\phi(\xi) = P(\xi)^{-1}$ where $P(\xi)$ is a homogeneous elliptic polynomial of sufficiently high degree and $\phi(\xi) = g(|\xi|^2)$ where $g(\zeta)$ is a is a univariate function which is analytic away from the origin and decays sufficiently fast at ∞ on the real axis. For some details in the second case see [2,7].

Another direction in one could consider extensions is the replacement of the lattice Z^n by a more general discrete set X. The case when X is finite is essentially treated in [6,15,17]. The general infinite case is not so clear.

References

[1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic Press, New York, 1967.

[2] M. D. Buhmann and M. J. D. Powell, Radial basis function interpolation on an infinite regular grid, preprint 1988, 24p.

[3] C. K. Chui, *Multivariate Splines*, CBMS Vol. 54, SIAM, Philadelphia, 1988.

[4] C. deBoor, How small can one make the derivatives of an interpolating function?, *J. Approx. Theory* 13(1975), 105-116.

[5] C. de Boor, K. Hollig, and S. Riemenschneider, Bivariate cardinal interpolation on a three-direction mesh, *Ill. J. Math.*, 29 (1985), 533-566.

[6] J. Duchon, Splines minimizing rotation-invariant seminorms in Sobolev spaces, in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller eds., Springer-Verlag, 1977, 85-100.

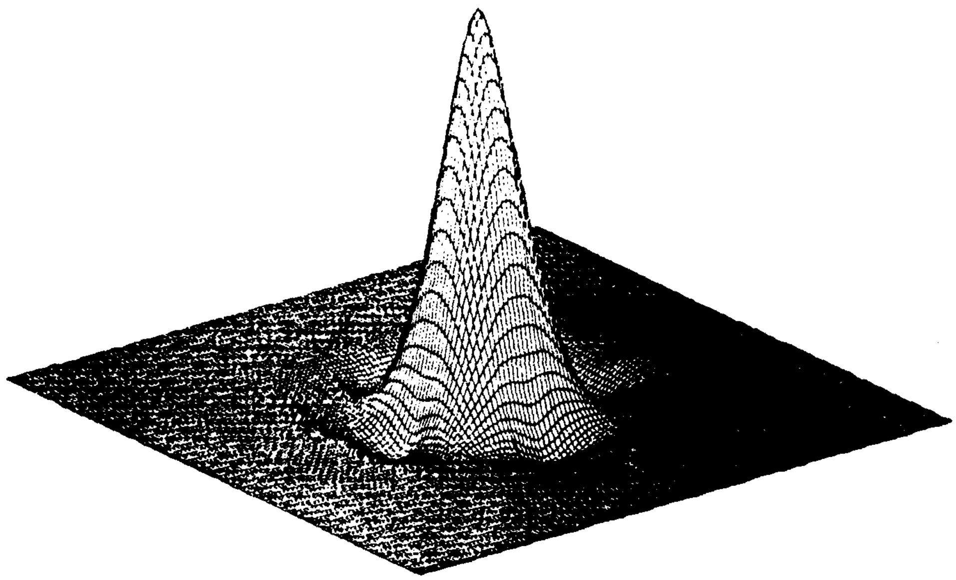

Figure 1: Plot of the surface $z = L_2(x, y)$.

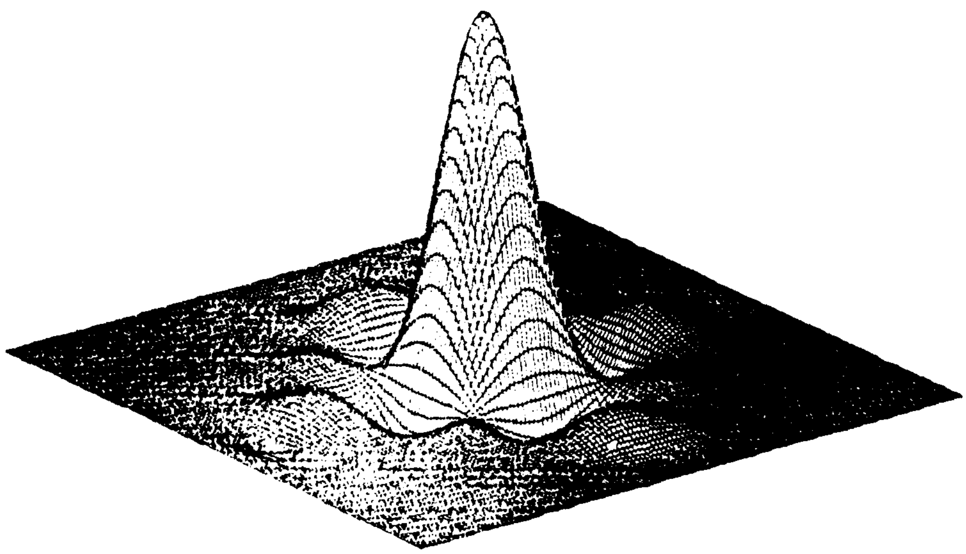

Figure 2: Plot of the surface $z = L_3(x, y)$.

[7] N. Dyn, Survey Lecture, to appear in *Approximation Theory VI*, C. Chui, L. Schumaker, J. Ward, eds., Academic Press, New York, 1989.

[8] M. Golomb and H. F. Weinberger, Optimal approximation and error bounds, in *On Numerical Approximation*, R. E. Langer, ed., Madison, 1959, 117-190.

[9] W. R. Madych, Polyharmonic splines, unpublished preprint 1978, 16p.

[10] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines, *Abstracts Amer. Math. Soc.* 7(1986), 378.

[11] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines, to appear in *J. Approx. Theory*.

[12] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines: a minimization property, technical report, BRC/MATH-TR-89-1.

[13] W. R. Madych and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions II, to appear in *Math. Comp.*

[14] W. R. Madych, Splines and entire functions, technical report, BRC/MATH-TR-89-2.

[15] J. Meinguet, Multivariate interpolation at arbitrary points made simple, *Z. Angew Math. Phys.*, 30(1979), 292-304.

[16] C. A. Micchelli, Interpolation of scattered data: distance matrices and conditionally positive definite functions. *Constr. Approx.*, 2(1986), 11-22.

[17] E. H. Potter, Multivariate Polyharmonic Spline Interpolation, Ph.D. thesis, Iowa State University, 1981.

[18] W. Schempp, *Complex Contour Integral Representation of Cardinal Spline Functions*, Contemporary Mathematics, Vol. 7, Amer. Math. Soc., Providence, 1982.

[19] I. J. Schoenberg, Contributions to the problem of approximation of equidistant data by analytic functions, *Quart. Appl. Math.*, 4 (1946), 45-99.

[20] I. J. Schoenberg, *Cardinal Spline Interpolation*, CBMS Vol. 12, SIAM, Philadelphia, 1973.

[21] G. Wahba, Spline bases, regularization, and generalized cross validation for solving approximation problems with large quantities of noisy data, *Approximation Theory III*, E. W. Cheney, ed., Academic Press, New York, 1980.

Hilbert Spaces for Estimators

W. R. Madych*
Department of Mathematics
University of Connecticut
Storrs, CT 06268.

Abstract

We outline a general procedure for reconstructing data from measurements based on a linear model. Roughly speaking, the method involves the construction of a Hilbert space on which the measurement functionals are continuous; the desired estimator is then simply the minimum norm solution in this space. Under certain conditions this construction results in a reproducing kernel Hilbert space. An application to seismic borehole tomography is included to illustrate the method and its uses.

1 Introduction.

In various circumstances one obtains data concerning a given quantity with the objective of determining certain features of that quantity. If one cannot measure the desired features directly then these features must be determined or approximated from indirect measurements. The case of indirect measurement raises a host of interesting questions and problems, some of which can be roughly summarized as follows:

- Do the measurements determine the desired features uniquely?
- If not, then how well can one approximate the desired features from the measurements?
- Give a constructive method for determining or approximating the desired features from the measurements.

It is the purpose of this note to address the third item on this list when the relationship between the desired quantity and the measurements is linear. The first two items are treated only incidentally.

*Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

Observe that these questions are essentially mathematical in nature and cannot be effectively addressed without a mathematical model relating the measurements to the desired features. The general setup is as follows:

The measurements are a collection of scalars, $f_1, \ldots, f_n$, which are functionals of a phantom f, the desired or unknown quantity. In particular if we call these functionals $\ell_1, \ldots, \ell_n$, then $f_1 = \ell_1(f), \ldots, f_n = \ell_n(f)$. The desired quantity may be the phantom f itself or certain features of f which may be modeled by other functionals of f. Of course in most instances one can only hope to reconstruct an approximation of f or such functionals. In what follows we will denote such an approximation of f by $\tilde{f}$.

In the case when the functionals $\ell_1, \ldots, \ell_n$ are linear and are continuous on some Hilbert space $\mathcal{H}$, which presumably contains the phantom, a popular and natural choice for $\tilde{f}$ is the orthogonal projection of f onto the subspace generated by these functionals. This projection can be computed from the given data by well known methods and has several interesting properties. Perhaps the most significant of these is the fact that the mapping $(f_1, \ldots, f_n) \rightarrow \tilde{f}$ is linear. Furthermore this projection is that unique element of minimal $\mathcal{H}$ norm which satisfies the data and the resulting estimator is optimal in a certain sense, see [4] and [8].

Unfortunately, in many circumstances, the linear functionals are not continuous on some obvious Hilbert space nor is the phantom necessarily a member of such a space. In this paper we address a fairly common situation of this type which may be described abstractly as follows.

The phantom f is an element of some ambient linear space $\mathcal{W}$ and the set $\{\ell_1, \ldots, \ell_n\}$ is a collection of linearly independent linear functionals defined on a linear subspace of $\mathcal{W}$ which contains f. The collection $\{\phi_j\}$, $j = 1, \ldots,$ is a complete orthonormal set in some Hilbert space $\mathcal{H}$ such that $\ell_k(\phi_j)$ is a well defined scalar for all j and k. Of course $\mathcal{H}$ is assumed to be a subspace of $\mathcal{W}$ but the linear functionals are not assumed to be continuous on $\mathcal{H}$. The values $\ell_1(f), \ldots, \ell_n(f)$ are known and the problem is to find an approximation $\tilde{f}$ of f.

In this case, motivated by the continuous example, we construct a Hilbert space $\mathcal{H}_1$ on which the linear functionals $\ell_1, \ldots, \ell_n$ are continuous and take the approximation $\tilde{f}$ of f to be the orthogonal projection of f onto the subspace generated by these functionals in $\mathcal{H}_1$. The details of this construction and its properties are given below in section 2.

We indicate the connection between our approach and the theory of reproducing kernel Hilbert space in the first subsection of the third and final section. An application of the method to a model of seismic borehole tomography is given in subsection 3.2.

We wish to thank Israel Koltracht who showed us the results of as yet unpublished and ongoing work concerning the null space of the matrix M mentioned in connection with the basis defined by equation (31) in subsection 3.2.

2 Details.

First we summarize the basic setup.

- $\mathcal{W}$ is a linear space which contains the phantom f.
- $\phi_1, \phi_2, \ldots,$ is a complete orthonormal set in a Hilbert space $\mathcal{H}$. $\mathcal{H}$ is a subspace of $\mathcal{W}$.
- $\ell_1, \ldots, \ell_n$ is a collection of linearly independent linear functionals such that $\ell_k(\phi_j)$ is a well defined scalar for each k and j. In particular, we assume that the n sequences $\{\ell_k(\phi_1), \ell_k(\phi_2), \ldots\}$ are linearly independent; furthermore, $\ell_j(f)$ is a well defined scalar for each j. In other words, f and the sequence $\phi_1, \phi_2, \ldots,$ are in the domain of each ℓ_j.

We emphasize that the functionals ℓ_j are not necessarily defined on all of $\mathcal{W}$ or $\mathcal{H}$ and of course they are not necessarily continous on $\mathcal{H}$. Also, f is not assumed to be in $\mathcal{H}$.

The problem we address below is the following:

Given the values $\ell_1(f), \ldots, \ell_n(f)$ find an estimator $\tilde{f}$ of f.

As is customary, symbols $\langle f, g\rangle$ denote the inner product of two elements, f and g, in $\mathcal{H}$ and $\hat{f}_j = \langle f, \phi_j\rangle$, $j = 1, 2, \ldots$. Thus $\{\hat{f}_1, \hat{f}_2, \ldots\}$ is the sequence of Fourier coefficients of f with respect to the complete orthonormal system $\phi_1, \phi_2, \ldots,$ and $\mathcal{H}$ may be regarded as a collection of such sequences via the standard identification. The norm $\|f\|$ of an element f of $\mathcal{H}$ is defined by $\|f\|^2 = \langle f, f\rangle$ of course. In the discussion below, we assume the field of scalars to be the complex numbers.

2.1 Simple interpolation.

Consider the matrix

$$M = (\ell_i(\phi_j)) \ , \quad i, j = 1, \ldots, n. \tag{1}$$

This matrix is not necessarily invertible in the general case. However, in view of the fact that the sequences $\{\ell_j(\phi_1), \ell_j(\phi_2), \ldots\}$, $j = 1, \ldots, n,$ are linearly independent, it should be clear that, by reordering the ϕ_j's if necessary, M has a non-zero determinant and hence is invertible.

In any case, if M is not singular, we can always find an element $\tilde{f}$ which satisfies

$$\ell_k(\tilde{f}) = \ell_k(f) \ , \ k = 1, \ldots, n, \tag{2}$$

by simply setting

$$\tilde{f} = \sum_{j=1}^{n} a_j \phi_j \tag{3}$$

where the a_j's are scalars which satisfy the system of equations

$$\sum_{j=1}^{n} a_j \ell_k(\phi_j) = \ell_k(f)\ , \ k = 1, \ldots, n. \tag{4}$$

Since M is invertible, a unique set of such a_j's exists.

Note that this method can be applied even if the basis elements $\phi_1, \ldots, \phi_n$ are not necessarily orthogonal as long as the corresponding matrix M given by (1) is not singular. The resulting estimator, $\tilde{f}$ is a simply a linear combination of $\phi_1, \ldots, \phi_n$ which interpolates the data $\ell_1(f), \ldots, \ell_n(f)$, namely, $\ell_1(\tilde{f}) = \ell_1(f), \ldots, \ell_n(\tilde{f}) = \ell_n(f)$. The properties of such a solution depend on the relationship between f, the ℓ's, and ϕ's and are not well documented in the general case.

In the development below we give an alternate method for computing an approximant $\tilde{f}$ in the general case.

2.2 A Hilbert Space Approach.

Let h_j, $j = 1, 2, \ldots$, be a sequence of non-negative real numbers such that

$$\sum_{j=1}^{\infty} h_j |\ell_k(\phi_j)|^2 < \infty \tag{5}$$

for all k. Set $w_j = h_j^{-1}$ and consider the subspace H_1 of $\mathcal{H}$ consisting of those elements f for which

$$||f||_1 = \left(\sum_{j=1}^{\infty} |\hat{f}_j|^2 w_j \right)^{1/2} \tag{6}$$

is finite. (In the case $h_j = 0$ the term $|\hat{f}_j|^2 w_j$ is taken to be zero if $\hat{f}_j$ is zero; otherwise it fails to be finite.) The Hilbert space $\mathcal{H}_1$ is the completion of H_1 in the norm $||f||_1$ defined by (6). The inner product in $\mathcal{H}_1$ is given by

$$\langle f, g \rangle_1 = \sum_{j=1}^{\infty} \hat{f}_j \overline{\hat{g}_j} w_j. \tag{7}$$

Proposition 1 *The linear functionals $\ell_1, \ldots, \ell_n$ are continuous on $\mathcal{H}_1$.*

Proof For f in $\mathcal{H}_1$ we may write

$$\begin{aligned} |\ell_k(f)| &= |\textstyle\sum_{j=1}^{\infty} \hat{f}_j \ell_k(\phi_j)| \\ &\leq \left(\textstyle\sum_{j=1}^{\infty} |\hat{f}_j|^2 w_j\right)^{1/2} \left(\textstyle\sum_{j=1}^{\infty} |\ell_k(\phi_j)|^2 h_j\right)^{1/2}. \end{aligned}$$

Since $c_k^2 = \sum_{j=1}^{\infty} |\ell_k(\phi_j)|^2 h_j$ is finite the last inequality may be re-expressed as

$$|\ell_k(f)| \leq c_k \|f\|_1$$

which implies continuity of ℓ_k. ■

Proposition 2 *If the sequence h_j, $j = 1, 2, \ldots$, is bounded then $\mathcal{H}_1$ is a subspace of $\mathcal{H}$. If, in addition, all the h_j's are non-zero, then $\mathcal{H}_1$ is dense in $\mathcal{H}$ and the set $\ell_1, \ldots, \ell_n$ is linearly independent in $\mathcal{H}_1$.*

Proof To see the first statement simply write

$$\|f\|^2 = \sum_{j=1}^{\infty} |\hat{f}_j|^2 \leq \max_j \{h_j\} \sum_{j=1}^{\infty} |\hat{f}_j|^2 w_j = \max_j \{h_j\} \|f\|_1^2 .$$

The second statement follows from the fact that the subspace consisting of those f's which have only a finite number of non-zero Fourier coefficients is dense in both $\mathcal{H}_1$ and $\mathcal{H}$. ■

Recall that if ℓ is a continuous linear functional on $\mathcal{H}_1$ then there is an element g_ℓ of $\mathcal{H}_1$ so that $\ell(f) = \langle f, g_\ell \rangle_1$. In fact we may identify ℓ with g_ℓ. Doing so, it is clear that

$$\ell_k = \sum_{j=1}^{\infty} h_j \overline{\ell_k(\phi_j)} \phi_j \tag{8}$$

for $k = 1, \ldots, n$. These series converge in $\mathcal{H}_1$ by virtue of (5). Furthermore, if the h's are bounded then these series also converge in $\mathcal{H}$. The scalar product of ℓ_k and ℓ_m in $\mathcal{H}_1$ is given by

$$\langle \ell_k, \ell_m \rangle_1 = \sum_{j=1}^{\infty} \overline{\ell_k(\phi_j)} \ell_m(\phi_j) h_j . \tag{9}$$

Note that the collection $\ell_1, \ldots, \ell_n$ is not necessarily linearly independent as a subset of $\mathcal{H}_1$; this depends on the choice of the h's. However the h's can always be chosen so that $\ell_1, \ldots, \ell_n$ is linearly independent and, in what follows, we always assume that this is the case. In particular the matrix

$$L = (\langle \ell_k, \ell_m \rangle_1) \tag{10}$$

whose elements are defined by (9) is not singular and is positive definite.

Let $s(f)$ be defined by the formula

$$s(f) = \sum_{k=1}^{n} a_k \ell_k \tag{11}$$

where the coefficients a_k are computed from the data $\ell_1(f), \ldots, \ell_n(f)$ by solving the system of linear equations

$$\sum_{k=1}^{n} a_k \langle \ell_k, \ell_m \rangle_1 = \ell_m(f) \ , \ m = 1, \ldots, n. \tag{12}$$

Since L is not singular there is a unique set of a_k's which satisfies (12) and thus $s(f)$ is well defined. Observe that

$$\ell_m(s(f)) = \langle s(f), \ell_m \rangle = \ell_m(f) \tag{13}$$

for $m = 1, \ldots, n$.

In view of (13) we now have a method for constructing an estimator $\tilde{f}$ of f, namely $\tilde{f} = s(f)$. Furthermore, if we choose to regard f as an element of $\mathcal{H}_1$ then it is clear that $s(f)$ is simply the orthogonal projection of f onto S in $\mathcal{H}_1$. Here S denotes the subspace of $\mathcal{H}_1$ spanned by $\ell_1, \ldots, \ell_n$.

2.3 A biorthogonal representation.

A particularly elegant way of representing $s(f)$ is in terms of elements $\lambda_1, \ldots, \lambda_n$ which are biorthogonal to $\ell_1, \ldots, \ell_n$ in S.

Recall that a collection $\lambda_1, \ldots, \lambda_n$ is said to be normalized biorthogonal to $\ell_1, \ldots, \ell_n$ in S if it is a subset of S and

$$\langle \lambda_j, \ell_k \rangle_1 = \delta_{jk} \ , \ \ j, k = 1, \ldots, n \tag{14}$$

where δ_{jk} is the Kronecker delta. Here S, the subspace spanned by $\ell_1, \ldots, \ell_n$, is regarded as a closed subspace of $\mathcal{H}_1$ whose inner product coincides with that of $\mathcal{H}_1$.

If $\lambda_1, \ldots, \lambda_n$ is normalized biorthogonal to $\ell_1, \ldots, \ell_n$ in S then it easy to check that

$$\lambda_k = \sum_{j=1}^{n} b_{kj} \ell_j \ , \ \ k = 1, \ldots, n. \tag{15}$$

The coefficients b_{kj} can be found by solving

$$\sum_{j=1}^{n} b_{kj} \langle \ell_j, \ell_m \rangle_1 = \delta_{km} \ , \ \ k, m = 1, \ldots, n, \tag{16}$$

in other words, the matrix (b_{kj}) is simply L^{-1} where L is given by (10). In terms of such λ_k's $s(f)$ has the representation

$$s(f) = \sum_{k=1}^{n} \ell_k(f) \lambda_k. \tag{17}$$

It should be mentioned that in spite of the fact that representation (17) is elegant, in specific examples it may be awkward to use and other representations may be more convenient. This will be illustrated in the examples below.

2.4 Questions.

Now that we have a method for constructing an approximation $\tilde{f}$ of f, namely $\tilde{f} = s(f)$, two questions need attention: (i) What is the relationship between $\tilde{f}$ and f? (ii) What difficulties, if any, are there in computing $\tilde{f}$?

From the practical viewpoint the second question seems more significant than the first. For instance, if the approximant is essentially impossible to compute with any level of accuracy then questions involving degree of approximation are primarily of academic interest. On the other hand, it seems that the first question should not be completely ignored, even in the applied setting; certainly it must be of some significance to know relationships between the computed quantity and the desired solution.

In our setup the answer to the second question depends on the conditionedness of the matrix L defined by (10) and the ease and accuracy at which (11) can be evaluated. Any questions concerning computability depend on the specific choice of ℓ's, ϕ's, and h's and is very difficult to address in the general case. Furthermore, it should be mentioned that these can be controlled to some extent.

A quantity which is essential in addressing the first question is f which is often completely or partially unknown. On the other hand, it is possible to obtain error estimates in terms of certain nonlinear functionals of f. These nonlinear functionals include various norms of f and may vary from case to case. For a nice discussion concerning the type of information needed to obtain estimates of accuracy in many of the classical approximation problems see [4]. Below we give an estimate of the error $\|f - s(f)\|$ in terms of $\|f\|_1$ and other parameters associated with the ℓ's and ϕ's.

2.5 A general error estimate.

Recall that if f is in $\mathcal{H}_1$ then $s(f)$ is the orthogonal projection of f into S, the subspace spanned by $\ell_1, \ldots, \ell_n$ in $\mathcal{H}_1$. This implies that

$$\|f - s(f)\|_1^2 = \|f\|_1^2 - \|s(f)\|_1^2 \tag{18}$$

whenever f is regarded as an element of $\mathcal{H}_1$.

The following proposition gives an estimate of the error in the $\mathcal{H}$ norm.

Proposition 3 *Suppose the matrix M defined by (1) is invertible, the sequence $\{h_j\}$ is bounded, and f is in $\mathcal{H}_1$. Then*

$$\|f - s(f)\| \le \epsilon \|f\|_1 \tag{19}$$

where

$$\epsilon \le \max_{j \ge n+1} \{\sqrt{h_j}\} + \rho \left(\sum_{k=1}^{n} \sum_{j=n+1}^{\infty} |\ell_k(\phi_j)|^2 h_j \right)^{1/2}$$

and ρ is the spectral radius of M^{-1}.

Proof First observe that it suffices to prove the proposition in the case

(20) $$\ell_k(f) = 0 \ , \ \ k = 1, \ldots, n.$$

For if the result holds in this case then in the general case, since $\ell_k(f - s(f)) = 0$ for all k and $s(f - s(f)) = 0$, we may write

(21) $$\|f - s(f)\| \leq \epsilon \|f - s(f)\|_1.$$

Inequality (21) together with the fact that $\|f - s(f)\|_1 \leq \|f\|_1$ imply the desired result.

Now suppose that f satisfies (20). In other words

$$\sum_{j=1}^{\infty} \hat{f}_j \ell_k(\phi_j) = 0 \ , \ k = 1, \ldots, n$$

or

(22) $$\sum_{j=1}^{n} \ell_k(\phi_j)\hat{f}_j = - \sum_{j=n+1}^{\infty} \ell_k(\phi_j)\hat{f}_j \ , \ k = 1, \ldots, n.$$

If we denote the right hand side of (22) by y_k, set $y = (y_1, \ldots, y_n)^T$, and set $x = (\hat{f}_1, \ldots, \hat{f}_n)^T$ then (22) may be rewritten as

(23) $$Mx = y$$

From (23) it follows that

(24) $$|x|^2 = |M^{-1}y|^2 \leq \rho^2 |y|^2$$

where $|x|^2$ denotes the sum of the squares of the components of x and ρ is the spectral radius of M^{-1}. An application of the Schwartz inequality gives

$$|y|^2 \leq \sum_{k=1}^{n} \left\{ \left(\sum_{j=n+1}^{\infty} |\hat{f}_j|^2 h_j^{-1} \right) \left(\sum_{j=n+1}^{\infty} |\ell_k(\phi_j)|^2 h_j \right) \right\}$$

and since term in the first set of parentheses is dominated by $\|f\|_1^2$ we may write

(25) $$|y|^2 \leq \left(\sum_{k=1}^{n} \sum_{j=n+1}^{\infty} |\ell_k(\phi_j)|^2 h_j \right) \|f\|_1^2.$$

Finally

(26) $$\sum_{j=n+1}^{\infty} |\hat{f}_j|^2 \leq \max_{j \geq n+1} \{h_j\} \sum_{j=n+1}^{\infty} |\hat{f}_j|^2 h_j^{-1} \leq \max_{j \geq n+1} \{h_j\} \|f\|_1^2.$$

Since

(27) $$\|f\|^2 = \sum_{j=1}^{n} |\hat{f}_j|^2 + \sum_{j=n+1}^{\infty} |\hat{f}_j|^2$$

the desired result follows from (27) together with (24), (25), and (26). ■

3 Examples.

3.1 Point evaluation functionals.

Suppose that the ambient space $\mathcal{W}$ is a collection of functions defined on some set Ω. If x is an element of Ω then the point evaluation functional ℓ_x is defined by $\ell_x(f) = f(x)$ and generally makes sense for some, but not necessarily all, f in $\mathcal{W}$. In many such instances there are readily available collections of functions, $\phi_1, \phi_2, \ldots$, which are orthonormal with respect to appropriate scalar products $\langle \phi, \psi \rangle$ and for which $\ell_x(\phi_j)$ is well defined, for all j. Whenever this the case one can take $\mathcal{H}$ to be the Hilbert space generated by such a sequence. One can then follow the procedure outlined in subsection 2.2 to obtain a Hilbert space $\mathcal{H}_1$ on which ℓ_x is a continuous linear functional.

If $\mathcal{H}_1$ is such that ℓ_x is continuous for all x in Ω then it is a reproducing kernel Hilbert space. The theory of such spaces is well documented and has found wide application; a detailed account of the basic theory may be found in [2], examples and various applications may be found in [3], [4], [6], and the references cited there. In our setup the reproducing kernel K is given by

$$K(x, y) = \sum_{j=1}^{\infty} h_j \overline{\ell_x(\phi_j)} \ell_y(\phi_j) \,. \tag{28}$$

The variational theory of splines may be viewed in some sense as an application of the theory of reproducing kernel Hilbert spaces. Cardinal splines are a nice example of the use of the biorthogonal representation in this case, for instance see [1]. A comparison of the univariate polynomial B-splines with the corresponding cardinal splines also shows how other representations can be considerably more convenient in certain applications.

3.2 Seismic borehole tomography.

Borehole tomography plays a significant role in seismic image reconstuction. The linearized well to well model is particulary popular, for example, see [5] and [7]. The general setup may be described as follows:

The ambient space $\mathcal{W}$ of phantoms f can initially be taken to be the class of functions, $f(x, y)$, which are Lebesgue measurable on

$$\Omega = \{(x, y) \;:\; 0 \leq x \leq a,\; 0 \leq y \leq b\},$$

where a and b are fixed positive numbers. The linear functionals are averages along straigth lines between points on opposite vertical boundaries of Ω.

More precisely, let

$$\begin{aligned} x(t) &= ut \\ y(t) &= vt + y_0 \end{aligned} \tag{29}$$

be the parametric representation of the line from $(0, y_0)$ to (a, y_1); here $c = y_1 - y_0$,

$$r = \sqrt{a^2 + c^2},$$

$u = a/r$, $v = c/r$, and t is the arclength parameter, $0 \le t \le r$. The corresponding linear functional ℓ, evaluated at a feasable phantom f, is defined by

$$\ell(f) = \int_0^r f(x(t), y(t))dt \tag{30}$$

where $x(t)$ and $y(t)$ are given by (29) and the integration is taken with respect to the one dimensional Lebesgue measure dt.

Note that ℓ is not defined on all of $\mathcal{W}$, namely, $\ell(f)$ is defined only for those phantoms f for which the right hand side of (30) makes sense. Indeed, if we take $\mathcal{H} = L^2(\Omega)$, the class of function which are square integrable with respect to Lebesgue measure on Ω and with the usual scalar product, then $\ell(f)$ is not defined even for all such f.

The mathematical problem of well to well tomography then is to approximate f from a collection of averages $\ell(f)$ corresponding to different pairs of endpoints $(0, y_0)$ and (a, y_1).

Physically the quantity $\ell(f)$ represents the travel time of a signal, which may be acoustic, between the coordinates $(0, y_0)$ and (a, y_1). The phantom f represents the 'slowness' profile of a cross section of the earth, represented by Ω, between two bore holes, represented by the vertical boundary lines of Ω. For more details concerning the physical interpretation of this setup see the references cited above; in particular [7] gives several examples of different geometries used for data collection.

For the sake of definiteness we confine our attention to a specific geometry for data collection. The reader should have no difficulty in making the appropriate adjustments for other typical geometries.

The geometry we will consider is the following: Take

$$z_i = \frac{i - \frac{1}{2}}{m} b \ , \ i = 1, ...m,$$

and let ℓ_{ij} be the linear functional defined by (30) with $y_0 = z_i$ and $y_1 = z_j$, $i, j = 1, \ldots, m$. (It should be clear that in this case it is natural to index the functionals with multi-indicies.) The problem then is to find an estimator, $\tilde{f}$, using the data $\ell_{ij}(f)$, $i, j = 1, \ldots, m$. In the notation of section 2 note that $n = m^2$ and the values of n functionals of the phantom f are known. Clearly these functionals are linearly independent.

It is well known that this problem is notoriously ill-posed. For example, it is quite apparent that any two phantoms whose difference is a function g which depends only on the x variable and is odd on the interval $0 \le x \le a$, (namely, $g(a - x) = -g(x)$) cannot be distiquished from the data.

Nevertheless, because of the significance of this problem, various solution techniques are employed to obtain estimators. A popular technique is the so-called series expansion method; this is essentially simple interpolation described in subsection 2.2. The most common choice of basis functions seems to be the set ϕ_{ij}, $i, j = 1, \ldots, m$, also doubly indexed for convenience and defined by

$$(31) \qquad \phi_{ij}(x,y) = \begin{cases} 1 & \text{if } \frac{i-1}{m}a < x \leq \frac{i}{m}a \text{ and } \frac{j-1}{m}b < y \leq \frac{j}{m}b \\ 0 & \text{otherwise.} \end{cases}$$

The reason for this choice appears to be the pixel-like nature of such a basis. Unfortunately the corresponding matrix M, see formulas (1), (3), and (4) in subsection 2.1, is not invertible. For example, in the case $m = 20$ the number of degrees of freedom is $n = 400$ but the corresponding 400×400 matrix M has only rank 344. However the least squares solution of minimun norm of (4) can be found and the resulting estimator (3) calculated. Apparently such estimators seem to give acceptable results in many instances.

On the other hand, since the functionals ℓ_{ij} are linearly independent, the method of subsection 2.2 will at the very least result in estimators all of whose n degrees of freedom are constrained by the data.

A fairly natural choice of $\mathcal{H}$ in this case seems to be $\mathcal{H} = L^2(\Omega)$. There are many well known complete orthonormal sequences in this space which one can utilize to apply the procedure of subsection 2.2. For example, a sequence on which one can easily evaluate the functionals under consideration in closed form is given by

$$(32) \qquad \phi_{kl}(x,y) = \frac{1}{\sqrt{ab}} \psi\left(\frac{x}{a}\right) \psi\left(\frac{y}{b}\right) \ , \ k, l = 0, 1, \ldots$$

where

$$\psi_0(z) = 1 \text{ and } \psi_k(z) = \sqrt{2}\cos(k\pi z), \ k = 1, \ldots .$$

In the case both k and l are greater than zero ϕ_{kl} may be expressed as

$$\phi_{kl}(x,y) = \frac{1}{\sqrt{ab}} \left[\cos\pi\left(\frac{k}{a}x + \frac{l}{b}y\right) + \cos\pi\left(\frac{k}{a}x - \frac{l}{b}y\right) \right] ,$$

which is useful for evaluating (30) in closed form. Since the final expressions are rather cumbersome and calculation tedious but straightforward, we leave the explicit evaluation of (30) as an exercise for the interested reader.

Possible choices of h_{kl}'s which will guarantee that the series

$$\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} |\ell_{ij}(\phi_{kl})|^2 h_{kl}$$

converges and the corresponding matrix L defined in subsection 2.2 is not singular are

$$(33) \qquad h_{kl} = (1 + k^2 + l^2)^{-1-\epsilon}$$

where ϵ is a positive constant and

$$h_{kl} = \begin{cases} 1 & \text{if } 0 \le k \le n^2 \text{ and } 0 \le l \le n^2 \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

We emphasis that in addition to the ϕ's given by (32) there are many other possibilities. Currently I. Koltracht and the author are conducting numerical experiments in order to obtain quantitative comparisons of the estimators arising from this procedure with those arising from conventional methods. The results of these experiments will be reported on elsewhere.

References

[1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic, New York, 1967.

[2] N. Aronszajn, Theory of Reproducing Kernels, *Trans. Amer. Math. Soc.*, 68 (1950), 337-404.

[3] P. J. Davis, *Interpolation and Approximation*, Dover, New York, 1975.

[4] M. Golomb and H. F. Weinberger, Optimal Approximation and Error Bounds, in *On Numerical Approximation*, R. E. Langer, ed. Madison, 1959, 117-190.

[5] S. Ivanson, Seismic Borehole Tomography - Theory and Computational Methods, *Proc. IEEE*, 74 (1986), 328-338.

[6] J. Meinguet, Surface spline interpolation: basic theory and computational aspects, in *Approximation Theory and Spline Functions*, S. P. Singh, J. Burry, and B. Watson, eds., Reidel, Dordrecht, 1984, 124-142.

[7] J. E. Peterson, B. N. P. Paulson, and T. V. McEvilly, Application of algebraic reconstruction techniques to crosshole seismic data, *Geophysics*, 50 (1985), 1566-1580.

[8] J. F. Traub and H. Wozniakowski, *A General Theory of Optimal Algorithms*, Academic, New York, 1980.

Translation Invariant Multiscale Analysis

W. R. Madych*

Abstract

The notion of multiscale analysis introduced by R. R. Coifman and Y. Meyer is considered and the translation invariant case is characterized.

1 Introduction

Recall that a *dyadic multiscale analysis* of $L^2(R^n)$ is an increasing sequence $\mathcal{V} = \{V_j : j = \ldots, -1, 0, 1, 2, \ldots\}$ of closed subspaces of $L^2(R^n)$ which has the following properties:

(1). $\bigcup_{j=-\infty}^{\infty} V_j$ is dense in $L^2(R^n)$ and $\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$.

(2). $f(x)$ is in V_j if and only if $f(2x)$ is in V_{j+1}.

(3). There is a lattice Γ in R^n such that for every f in V_0 and every γ in Γ the function f_γ is in V_0. Here and in what follows we use the notation $f_\gamma(x) = f(x - \gamma)$.

(4). There are two positive constants $C_2 \geq C_1 > 0$ and a function g in V_0 such that V_0 is the closed linear span of g_γ, $\gamma \in \Gamma$, and

$$C_1^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2 \leq \int_{R^n} \sum_{\gamma \in \Gamma} |a_\gamma g_\gamma(x)|^2 dx \leq C_2^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2 .$$

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Preliminary report of work partially supported by a grant from the Air Force Office of Scientific Research.

An introduction to the subject may be found in [1,2]. A basic property of a multiscale analysis $\mathcal{V}$ is the following:

(5). There is a function ϕ in V_0 such that the collection $\{\phi_\gamma\}_{\gamma\in\Gamma}$ is an orthonormal basis in V_0.

This fact may be regarded as a substitute for (4) and plays an important role in what follows.

A dyadic multiscale analysis is ***translation invariant*** if all the translates of f, $\{f_y : y \in R^n\}$, are in V_0 whenever f is in V_0.

The canonical example of a translation invariant multiscale analysis of $L^2(R)$ is when V_0 is the collection of those functions in $L^2(R)$ whose Fourier transforms are supported in the interval $[-\pi, \pi]$. A natural choice of ϕ in this case is given by

$$\phi(x) = \frac{\sin \pi x}{\pi x}\,.$$

The point of this paper is to give a characterization of translation invariant multiscale analyses. For the sake of clarity in what follows we will restrict our attention to the case $n = 1$ and $\Gamma = Z$, the lattice of integers. The statements and arguments in the general case are completely analogous to this basic case.

We now briefly digress to list some of the conventions which are used here: The Fourier transform $\hat{f}$ of a function f is defined by

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} f(x)dx$$

whenever it makes sense and distributionally otherwise. Basic facts concerning Fourier transforms and distributions will be used without further elaboration in what follows. To avoid the pedantic repetition of "almost everywhere" and other modifying phrases which are inevitably necessary when dealing with functions defined almost everywhere, all equalities between functions and other related notions are interpreted in the distributional sense whenever possible. The term support is also used in the distributional sense; in particular the support of a function f in $L^2(R)$ is a well defined closed set. If W is a collection of tempered distributions then $\widehat{W}$ is the collection of Fourier transforms of elements of W, in other words $\widehat{W} = \{f : f = \hat{g} \text{ for some } g \text{ in } W\}$. For a subset Ω of R and a real number

r the sets $r\Omega$ and $\Omega + r$ are defined by $r\Omega = \{x : x = r\omega \text{ for some } \omega \text{ in } \Omega\}$ and $\Omega + r = \{x : x = \omega + r \text{ for some } \omega \text{ in } \Omega\}$; $L^2(\Omega)$ is the L^2 closure of the subspace of those functions in $L^2(R)$ whose support is contained in Ω. For notational simplicity we use Q to denote the closed interval $[-\pi, \pi]$.

We can now conveniently state our main observation.

Theorem *Suppose V is a translation invariant dyadic multiscale analysis of $L^2(R)$. Then $\hat{V}_0 = L^2(\Omega)$ where Ω is a closed subset of R which has the following properties:*

(a). $\Omega \subset 2\Omega$.

(b). $\{\Omega + 2\pi j\} \bigcap \{\Omega + 2\pi k\}$ *is a set of Lebesgue measure 0 for any pair of integers such that* $j \neq k$.

(c). $\bigcup_{k=-\infty}^{\infty} \{\Omega + 2\pi k\} = R$.

(d). $\bigcup_{k=1}^{\infty} L^2(2^k\Omega)$ *is dense in* $L^2(R)$.

Conversely, if V_k, $k \in Z$ is defined by $\hat{V}_k = L^2(2^k\Omega)$ where Ω is a closed subset of R which satisfies the properties above then the sequence of subspaces $\{V_k\}$ is a translation invariant multiscale analysis of $L^2(R)$.

Remark 1 In view of the example given above it is very tempting to conjecture that the set Ω in the Theorem must be of the form $\Omega = Q + \alpha$ for some real number α which satisfies $\pi < \alpha < \pi$. Certainly such Ω's satisfy the desired conditions. However the conditions of the Theorem are satisfied by Ω's which need not be connected as the following example due to Rudi Lorentz shows.

$$\Omega = \left[-\frac{5\pi}{4}, -\pi\right] \bigcup \left[-\frac{3\pi}{4}, \frac{3\pi}{4}\right] \bigcup \left[\pi, \frac{5\pi}{4}\right] .$$

Remark 2 Consider

$$\begin{aligned} \Omega_a &= [-1, 1] \bigcup [2\pi + 1, 4\pi - 1] \\ \Omega_b &= [-5, 5] \\ \Omega_c &= [-1, 1] \\ \Omega_d &= [0, 2\pi] . \end{aligned}$$

It is not difficult to verify that each Ω_α listed above is a closed set which fails to satisfy condition (α) but satisfies the remaining conditions in the Theorem. These examples show that conditions *(a)-(d)* are not redundant.

Remark 3 Note that condition *(a)* implies that 0 is contained in Ω. In addition to this it is clear that if Ω contains a neighborhood of the origin then it satisfies condition *(d)*. In view of this it seems reasonable to suspect that subsets Ω which satisfy the conditions of the Theorem must contain an open neighborhood of the origin. That this is not the case can be seen by considering the following example of Ω:

$$\{\bigcup_{k=1}^{\infty}[-(2-2^{-k})2^{-k}\pi, -2^{-k}\pi]\}\bigcup[0,\pi]\{\bigcup_{k=1}^{\infty}[(2-(2-2^{-k})2^{-k})\pi, (2-2^{-k})\pi]\} .$$

Remark 4 In view of the examples listed above it may be of some interest to obtain a significantly more lucid description of the set Ω than that given in the Theorem.

A corollary concerning wavelets generated by $\mathcal{V}$ is recorded at the end of Section 2.

2 Details

We begin by establishing a basic lemma. First recall that the indicator function of a set Ω is usually denoted by χ_Ω and satisfies

$$\chi_\Omega(\xi) = \begin{cases} 1 \text{ if } \xi \in \Omega \\ 0 \text{ otherwise.} \end{cases}$$

Lemma *Suppose $\mathcal{V}$ is a translation invariant dyadic multiscale analysis of $L^2(R)$ and ϕ is a function whose existence is guaranteed by (5). Then*

$$|\hat{\phi}| = \chi_\Omega$$

where χ_Ω is the indicator function of a closed set Ω which has properties (a)-(d) in the statement of the Theorem.

Let Ω be the support of $\hat{\phi}$. To prove the lemma we will first show that Ω satisfies property *(b)*.

Recall that (5) implies that for all f in V_0 we may write

$$\hat{f}(\xi) = g(\xi)\hat{\phi}(\xi) \tag{6}$$

where g is 2π periodic and square integrable over Q. In particular, since V_0 is translation invariant, ϕ_y is in V_0 so setting $\alpha = -y$ we may write

$$e^{i\alpha\xi}\hat{\phi}(\xi) = g(\xi)\hat{\phi}(\xi)$$

for some such g. Hence

$$e^{i\alpha(\xi-2\pi m)}\hat{\phi}(\xi - 2\pi m) = g(\xi - 2\pi m)\hat{\phi}(\xi - 2\pi m) = g(\xi)\hat{\phi}(\xi - 2\pi m)$$

which implies that

$$e^{i\alpha(\xi-2\pi m)} = g(\xi)$$

on $\Omega + 2\pi m$. For two different values of m the last equality implies that

$$e^{i\alpha(\xi-2\pi j)} = e^{i\alpha(\xi-2\pi k)}$$

on $\{\Omega + 2\pi j\}\bigcap\{\Omega + 2\pi k\}$. Re-expressing the last relation as

$$e^{i\alpha(\xi-2\pi k)}(e^{i2\pi\alpha(k-j)} - 1) = 0$$

it is clear that either α is an integer, j is equal to k, or $\{\Omega+2\pi j\}\bigcap\{\Omega+2\pi k\}$ is a set of measure zero. Since α may be any real number we conclude that $\{\Omega + 2\pi j\}\bigcap\{\Omega + 2\pi k\}$ has measure zero whenever $j \neq k$.

Now, since $\hat{\phi}_k(\xi) = e^{-ik\xi}\hat{\phi}(\xi)$, $k \in Z$, are orthonormal, we may write

$$\int_R \phi_k(x)\overline{\phi_\ell(x)}dx = \int_R e^{im\xi}|\hat{\phi}(\xi)|^2 d\xi = \tag{7}$$

$$\int_Q e^{im\xi}\sum_{j\in Z}|\hat{\phi}(\xi - 2\pi j)|^2 dx = \begin{cases} 1 \text{ when } m = 0 \\ 0 \text{ otherwise} \end{cases}$$

where $m = \ell - k$. The last equality implies that

$$\sum_{j\in Z}|\hat{\phi}(\xi - 2\pi j)|^2 = \frac{1}{2\pi}$$

on R and since $\{\Omega + 2\pi j\} \bigcap \{\Omega + 2\pi k\}$ has measure zero whenever $j \neq k$ we may conclude that

$$|\hat{\phi}(\xi)| = \frac{1}{\sqrt{2\pi}} \chi_\Omega(\xi)$$

and

$$\bigcup_{k \in Z} \{\Omega + 2\pi k\} = R \, .$$

To see (a) observe that (2) and the facts demonstrated above imply that

$$\chi_\Omega(\xi) = h(\xi)\chi_\Omega(\xi/2)$$

where h is 4π periodic and square integrable over $2Q$. Since $\chi_{2\Omega}(\xi) = \chi_\Omega(\xi/2)$, the last equality involving h implies that χ_Ω vanishes whenever $\chi_{2\Omega}$ does so $\Omega \subset 2\Omega$.

Finally, the fact that $\bigcup_{k=1}^{\infty} L^2(2^k\Omega)$ is dense in $L^2(R)$ is an immediate consequence of property (1). The proof of the Lemma is complete.

Now, suppose ϕ and Ω are as in the Lemma and its proof. Since V_0 consists of functions f which satisfy (6) it is clear that $\widehat{V}_0$ is contained in $L^2(\Omega)$.

To see that $\widehat{V}_0 = L^2(\Omega)$ let f be any element in $L^2(\Omega)$ and let h be defined by

$$h(\xi) = \begin{cases} \overline{\hat{\phi}(\xi)}/|\hat{\phi}(\xi)| & \text{if } \hat{\phi}(\xi) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

By virtue of the properties of Ω established above it is clear that

$$\frac{1}{\sqrt{2\pi}} \chi_\Omega(\xi) = \left(\sum_{j \in Z} h(\xi - 2\pi j) \right) \hat{\phi}(\xi)$$

and

$$\hat{f}(\xi) = g(\xi)\hat{\phi}(\xi)$$

where

$$g(\xi) = \sqrt{2\pi} \left(\sum_{j \in Z} \hat{f}(\xi - 2\pi j) \right) \left(\sum_{j \in Z} h(\xi - 2\pi j) \right) .$$

Thus f satisfies (6) and hence we may conclude that $L^2(\Omega)$ is contained in $\widehat{V}_0$.

The Lemma together with the last observation imply the first assertion of the Theorem.

To see the converse, let $\mathcal{V}$ be the sequence of subspaces $\{V_k\}$, $k \in Z$, defined by $\widehat{V}_k = L^2(2^k\Omega)$ where Ω is a closed set which satisfies properties *(a)-(d)* of the Theorem.

The fact that $\mathcal{V}$ is translation invariant and, in particular, satisfies property (3) with $\Gamma = Z$ is an immediate consequence of the definition. Property (2) is also immediate. That $\mathcal{V}$ is an increasing sequence of subspaces and $\bigcup_{k\in Z} V_k$ is dense in $L^2(R)$ are consequences of properties *(a)* and *(d)*.

That $\bigcap_{k\in Z} V_k = \{0\}$ follows from the fact that the measure of Ω is finite. Indeed, its measure is 2π which can be seen from

$$\int_R \chi_\Omega(\xi)d\xi = \sum_{j\in Z}\int_{Q+2\pi j}\chi_\Omega(\xi)d\xi = \int_Q \sum_{j\in Z}\chi_\Omega(\xi - 2\pi j)d\xi = 2\pi$$

by using properties *(b)* and *(c)*.

Finally, to see property (5) take

$$\hat{\phi} = \frac{1}{\sqrt{2\pi}}\chi_\Omega$$

and use properties (b) and (c) to write (7) which shows that $\phi_k(x)$, $k \in Z$, are orthonormal and, for f in V_0,

$$\hat{f}(\xi) = \sqrt{2\pi}\left(\sum_{k\in Z}\hat{f}(\xi - 2\pi k)\right)\hat{\phi}(\xi)$$

or

$$f(x) = \sum_{k\in Z} f(j)\phi(x-k)$$

which shows that they are complete in V_0.

This completes the proof of the Theorem.

Remark 5 Suppose $\mathcal{V}$ is a translation invariant multiscale analysis and Ω is a closed set such that $\widehat{V}_0 = L^2(\Omega)$. Then if W_0 is the orthogonal complement of V_0 in V_1, $\widehat{W}_0 = L^2(\Upsilon)$ where $\Upsilon = 2\Omega \setminus \Omega$. Let ψ be such that the set $\{\psi_k : k \in Z\}$ is an orthonormal basis for W_0. Such a ψ may be referred to

as a wavelet. Using reasoning analogous to the proof of the Theorem, it is clear that ψ is a wavelet if and only if

$$|\hat{\psi}| = \frac{1}{\sqrt{2\pi}}\chi_{\Upsilon} \,.$$

Now, an analyzing wavelet in the sense of Meyer, [1], is globally integrable and hence its Fourier Transform must be continuous. Clearly ψ is not such an analyzing wavelet.

Corollary *A translation invariant multiscale analysis cannot give rise to analyzing wavelets in the sense of Meyer.*

References

[1] J. M. Combs, A. Grossman, and Ph. Tchamitchian, eds., *Wavelets*, Proceedings of the International Conference in Marseille, 1987, Springer-Verlag, Berlin, 1989.

[2] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, 41 (1988), 909-996.